Н. Вирт

Построение компиляторов

Москва
«Книга по Требованию»

УДК 004.42
ББК 32.811
В52

Вирт Н.
В52 Построение компиляторов / Н. Вирт – М.: Lennex Corp, — Подготовка макета: ООО «Книга по Требованию», 2023. – 186 с.

ISBN 978-5-94074-585-3

Книга известного специалиста в области информатики Никлауса Вирта написана по материалам его лекций по вводному курсу проектирования компиляторов. На примере простого языка Оберон-0 рассмотрены все элементы транслятора, включая оптимизацию и генерацию кода. Приведен полный текст компилятора на языке программирования Оберон. Для программистов, преподавателей и студентов, изучающих системное программирование и методы трансляции. Прилагается компакт-диск с полным комплектом примеров программ в системе программирования Блэкбокс.

ISBN 978-5-94074-585-3

© Lennex Corp, 2023
© Н. Вирт, 2023

Никлаус Вирт

Построение компиляторов

ДМК Пресс
издательство
Москва, 2010

УДК 32.973.26-018.2
ББК 004.438
В52

Никлаус Вирт
Построение компиляторов / Пер. с англ. Борисов Е. В., Чернышов Л. Н. – М.: ДМК Пресс, 2010. – 192 с.: ил.

ISBN 978-5-94074-585-3

Книга известного специалиста в области информатики Никлауса Вирта написана по материалам его лекций по вводному курсу проектирования компиляторов. На примере простого языка Оберон-0 рассмотрены все элементы транслятора, включая оптимизацию и генерацию кода. Приведен полный текст компилятора на языке программирования Оберон.

Для программистов, преподавателей и студентов, изучающих системное программирование и методы трансляции.

Содержание компакт-диска:

Базовая конфигурация системы Блэкбокс с коллекцией модулей, реализующих оригинальный компилятор с языка Оберон-0 и компилятор, адаптированный под Блэкбокс.
Базовые инструкции по работе в системе Блэкбокс.
Полный перевод документации системы Блэкбокс на русский язык.
Конфигурация системы Блэкбокс для использования во вводных курсах программирования в университетах.
Конфигурация системы Блэкбокс для использования в школах (полная русификация меню, сообщений компилятора, с возможностью использования ключевых слов на русском и других национальных языках).
Доклады участников проекта Информатика-21 по опыту использования системы Блэкбокс в обучении программированию.
Оригинальные дистрибутивы системы Блэкбокс 1.5 (основной рабочий) и 1.6rc6.
Инструкции по работе в Блэкбоксе под Linux/Wine.
Дистрибутив оптимизирующего компилятора XDS Oberon (версии Linux и MS Windows).
OberonScript – аналог JavaScript для использования в Web-приложениях.

This is a slightly revised version of the book published by Addison-Wesley in 1996

ISBN 0-201-40353-6 (анг.)
© N. Wirth, 1985 (Oberon version: August 2004)
© Перевод с английского Борисов Е. В., Чернышов Л. Н., 2010
ISBN 978-5-94074-585-3
© Оформление, издание, ДМК Пресс, 2010

Краткое содержание

ОТ АВТОРОВ ПЕРЕВОДА 10

ВВЕДЕНИЕ 12

ГЛАВА 1. ВВЕДЕНИЕ 15

ГЛАВА 2. ЯЗЫК И СИНТАКСИС 19

ГЛАВА 3. РЕГУЛЯРНЫЕ ЯЗЫКИ 27

ГЛАВА 4. АНАЛИЗ КОНТЕКСТНО-СВОБОДНЫХ ЯЗЫКОВ 33

ГЛАВА 5. АТРИБУТНЫЕ ГРАММАТИКИ И СЕМАНТИКИ 45

ГЛАВА 6. ЯЗЫК ПРОГРАММИРОВАНИЯ ОБЕРОН-0 51

ГЛАВА 7. СИНТАКСИЧЕСКИЙ АНАЛИЗАТОР ДЛЯ ОБЕРОНА-0 55

ГЛАВА 8. УЧЕТ КОНТЕКСТА, ЗАДАННОГО ОБЪЯВЛЕНИЯМИ 65

ГЛАВА 9. RISC-АРХИТЕКТУРА КАК ЦЕЛЬ 75

ГЛАВА 10. ВЫРАЖЕНИЯ И ПРИСВАИВАНИЯ 81

ГЛАВА 11. УСЛОВНЫЕ И ЦИКЛИЧЕСКИЕ ОПЕРАТОРЫ И ЛОГИЧЕСКИЕ ВЫРАЖЕНИЯ 95

ГЛАВА 12. ПРОЦЕДУРЫ И КОНЦЕПЦИЯ ЛОКАЛИЗАЦИИ 109

ГЛАВА 13. ЭЛЕМЕНТАРНЫЕ ТИПЫ ДАННЫх 125

ГЛАВА 14. ОТКРЫТЫЕ МАССИВЫ, УКАЗАТЕЛЬНЫЙ И ПРОЦЕДУРНЫЙ ТИПы 131

ГЛАВА 15. МОДУЛИ И РАЗДЕЛЬНАЯ КОМПИЛЯЦИЯ 141

ГЛАВА 16. ОПТИМИЗАЦИЯ И СТРУКТУРА ПРЕ/ПОСТПРОЦЕССОРА 153

ПРИЛОЖЕНИЕ А. СИНТАКСИС 164

ПРИЛОЖЕНИЕ В. НАБОР СИМВОЛОВ ASCII 167

ПРИЛОЖЕНИЕ С. КОМПИЛЯТОР ОБЕРОН-0 168

ЛИТЕРАТУРА 191

Содержание

От авторов перевода 10
О книге 10
О переводе 10

Введение 12
Предисловие 12
Благодарности 14

Глава 1. Введение 15

Глава 2. Язык и синтаксис 19
2.1. Упражнения 24

Глава 3. Регулярные языки 27
3.1. Упражнение 32

Глава 4. Анализ контекстно-свободных языков 33
4.1. Метод рекурсивного спуска 34
4.2. Таблично-управляемый нисходящий синтаксический анализ 38
4.3. Восходящий синтаксический анализ 40
4.4. Упражнения 42

Глава 5. Атрибутные грамматики и семантики 45
5.1. Правила типов 46
5.2. Правила вычислений 47
5.3. Правила трансляции 48
5.4. Упражнение 49

Глава 6. Язык программирования Оберон-0 51
6.1. Упражнение 54

Глава 7. Синтаксический анализатор для Оберона-0 55
7.1. Лексический анализатор 56
7.2. Синтаксический анализатор 57
7.3. Устранение синтаксических ошибок 59
7.4. Упражнения 64

Глава 8. Учет контекста, заданного объявлениями 65
8.1. Объявления 66
8.2. Записи о типах данных 68
8.3. Представление данных во время выполнения 69
8.4. Упражнения 73

Глава 9. RISC-архитектура как цель 75
9.1. Ресурсы и регистры 76

Глава 10. Выражения и присваивания 81
10.1. Прямая генерация кода по принципу стека 82
10.2. Отсроченная генерация кода 84
10.3. Индексированные переменные и поля записей 89
10.4. Упражнения 94

Глава 11. Условные и циклические операторы и логические выражения 95
11.1. Сравнения и переходы 96
11.2. Условные и циклические операторы 97
11.3. Логические операции 101
11.4. Присваивание логическим переменным 105
11.5. Упражнения 106

Глава 12. Процедуры и концепция локализации 109
12.1. Организация памяти во время выполнения 110
12.2. Адресация переменных 112
12.3. Параметры 114
12.4. Объявления и вызовы процедур 116

12.5. Стандартные процедуры 121
12.6. Процедуры-функции 122
12.7. Упражнения 123

Глава 13. Элементарные типы данных 125
13.1. Типы REAL и LONGREAL 126
13.2. Совместимость между числовыми типами данных 127
13.3. Тип данных SET 129
13.4. Упражнения 130

Глава 14. Открытые массивы, указательный и процедурный типы 131
14.1. Открытые массивы 132
14.2. Динамические структуры данных и указатели 133
14.3. Процедурные типы 136
14.4. Упражнения 138

Глава 15. Модули и раздельная компиляция 141
15.1. Принцип скрытия информации 142
15.2. Раздельная компиляция 143
15.3. Реализация символьных файлов 145
15.4. Адресация внешних объектов 149
15.5. Проверка конфигурационной совместимости 150
15.6. Упражнения 152

Глава 16. Оптимизация и структура пре/постпроцессора 153
16.1. Общие соображения 154
16.2. Простые оптимизации 155
16.3. Исключение повторных вычислений 156
16.4. Распределение регистров 157
16.5. Структура пре/постпроцессорного компилятора 158
16.6. Упражнения 162

Приложение А. Синтаксис 164
А1. Оберон-0 164
А2. Оберон 164
А3. Символьные файлы 166

Приложение B. Набор символов ASCII 167

Приложение C. Компилятор Оберон-0 168
C.1. Лексический анализатор 169
C.2. Синтаксический анализатор 172
C.3. Генератор кода 182

Литература 191

От авторов перевода

О книге

Давно известно, что лучший способ постичь секреты мастерства – это наблюдать за работой мастера. Эта небольшая, но насыщенная информацией книжка, по сути дела, представляет собой отчет о такой работе. Ну а то, что ее автор – настоящий мастер своего дела, сомнению не подлежит, потому что имя профессора Никлауса Вирта ни в каких дополнительных рекомендациях не нуждается. Эта книга – своего рода мастер-класс, который дает своим ученикам всемирно известный маэстро. Она не является ни «тяжелой» теоретической монографией, ни сборником наставлений и поучений увенчанного лаврами мэтра. Эта книжка – *практическое* пособие для всех тех любознательных людей, кто желает разобраться и понять, что такое компилятор и как он устроен. По мнению автора, без этого ни один программист не может называть себя квалифицированным специалистом.

В отличие от многочисленных книг, которые исчерпывающе описывают и теорию, и разнообразные методы синтаксического анализа, перевода и компиляции, эта книжка посвящена реализации одного-единственного компилятора *современного* языка программирования для *конкретного* компьютера. Но это нисколько не умаляет ее достоинства. Если обычные книги после прочтения почти всегда оставляют читателя наедине с вопросом «А что же дальше? Где же результат?» или с загадочными, полными опечаток текстами готовых программ, то эта небольшая книжка расставляет практически все точки над i, проводя читателя от *самого начала* до *самого конца* процесса разработки компилятора, попутно предупреждая его о неверных шагах и давая ему в руки богатый практический материал. Автор придерживается принципа «Делай со мной. Делай, как я. Делай лучше меня».

Таким образом, книга Н. Вирта – безусловно, не только прекрасное дополнение к многочисленным и столь же прекрасным фундаментальным трудам по этой теме, но может и должна использоваться в качестве практического пособия по изучению компиляторов. Кроме того, простота и доступность преподнесения довольно сложного материала снимает с него покров таинственности и делает его доступным практически каждому любителю программирования. Остается только сожалеть о том, что эта книга не была своевременно переведена и издана у нас.

Для практического использования текст компилятора Оберон-0, о котором идет речь в книге, адаптирован к системе БлэкБокс (BlackBox Component Builder – вариант системы Оберон). Оригинальные и адаптированные исходные тексты компилятора можно найти на сайте www.oberoncore.ru.

О переводе

Несколько слов о переводе.

В силу того, что мы имеем дело не с развернутой монографией, а с конспектом лекций, каждая фраза, часто облекаемая в форму тезиса, до предела насыщена

информацией. Поэтому наша основная задача при переводе состояла в том, чтобы сохранить лаконичность и информационную насыщенность авторского текста и при этом максимально точно довести его суть до читателя, не поддаваясь искушению сдобрить его отсебятиной.

Несмотря на царящие до сих пор «разброд и шатания» в терминологии по этой теме, мы при переводе, следуя за автором, отдавали предпочтение наиболее устоявшимся, хотя и не всегда правильным и точным, терминам.

В связи с этим нельзя не упомянуть о терминах «front-end» и «back-end». Они уже давно употребляются в разнообразной англоязычной технической литературе, но тем не менее до сих пор не находят адекватных русскоязычных эквивалентов. Чаще всего их перевод зависит от контекста. Применительно к компиляторам наиболее точными их русскими аналогами являются, пожалуй, «машинно-независимая часть» и «машинно-зависимая часть» соответственно. Однако мы, теперь уже следуя авторской лаконичности, предпочли им более абстрактные и менее точные, но более короткие термины – «препроцессор» и «постпроцессор» соответственно.

Кроме того, список литературы пронумерован, и именные ссылки на него в тексте заменены номерными. К списку литературы добавлено несколько более поздних публикаций.

Авторы перевода выражают благодарность В. Н. Лукину за прочтение перевода и сделанные замечания.