

СИСТЕМА SIS79/GENERA

Пал Керекфи

1. СИСТЕМА РЕДАКТИРОВАНИЯ ПРОГРАММ ГЕНЕРА

1.1. Введение

Целью разработки программных средств обычно является создание процедур, способствующих облегчению и эффективности использования ЭВМ.

Генерирование /редактирование/ программ является методом помогающим программистам. Под генерированием подразумевается процедура, при помощи которой на базе команд /программ/, написанных на языке более высокого уровня, создается программа на языке более низкого уровня. /Этот язык имеет транслятор в данной программной конфигурации или встроенный в аппаратурные средства./ В таком смысле генерирующими программами являются трансляторы различных языков высокого уровня, а также предварительные трансляторы. Генерирование программ одновременно обеспечивает эффективность и более простое использование. Часто встречаем проблему, что программы, написанные в общей форме, недостаточно эффективные, и наоборот, программы, написанные с учетом эффективности, слишком сильно связаны с данной задачей и трудно их использовать даже при небольших изменениях параметров задачи. Поэтому в большин-

стве случаев дается компромисное решение, которое зависит от характера задачи, т.е. программа, которая более или менее общая и более или менее эффективная. В решении этих проблем поможет редактирование /генерирование/ программ.

Первой фазой проектирования системы является создание структуры эффективной целевой программы. Это является трудоемкой задачей и требует большое внимание. Написать процедуру генерирования уже значительно проще, особенно если она входит в систему, обеспечивающую подходящие условия для таких процедур. Система "Генера" была создана именно для такой цели, чтобы способствовать получению процедур редактирования.

1.2. Общее описание системы "Генера"

Целью системы "Генера" является включение процедур редактирования в единую систему. Упрощение и согласование работы программистов /системных программистов/, а также обеспечение единого, наглядного использования процедур компоновки. Система является модульной в большой степени и обеспечивает произвольное соединение процедур, и тем самым решение нескольких задач внутри одной, генерированной программы.

1.3. Команды

Произвольное соединение процедур редактирования в одну программу обеспечивается под управлением команд с произвольным содержанием /в нашем случае это программа на языке ФОРТРАН/. Команды находятся на входном файле. Строки, которые не содержат команд для "Генера", без изменения переносятся системой на выходной файл, а на место команд "Генера" записываются программные фрагменты, созданные процедурами редактирования. Команды управления начинаются специальным символом.

1.4. Построение системы "Генера"

Главными частями системы являются:

- Предварительная обработка
- Подготовка
- Обработка
- Окончание

Во время предварительной обработки программа производит формальную проверку команд и их параметров. В случае нахождения ошибки программа осведомляет об этом пользователя и после предварительной обработки программа останавливается. Во время проверки команд программа собирает данные о количестве памяти и файлов, необходимых для генерирования.

На основе этого в фазе подготовки проверяются необходимые для программы ресурсы.

Во время обработки программа переписывает содержание входного файла на выходной файл /генерированный программный файл/. Если при этом "Генера" находит команду /запись начинается с соответствующим специальным символом/, тогда прекратит переписывание и вызывает соответствующую подсистему для исполнения команды.

После полной обработки входного файла программа в соответствии с выбранным способом прогона обеспечивает трансляцию и прогон генерированной программы. Способ прогона /и другие параметры/ задаются пользователем с помощью команды OPTION.

Частью системы является подсистема макрогенератора. С помощью этого пользователь может использовать макро, определенные в системе, а также может определить новые.

2. КОМАНДЫ В SIS 79 /СИС 79/

В дальнейшем изложены команды языка СИС 79, связанные с системой Генера. СИС 79, как и другие системы построенные на Генера, легко расширяемый язык и может быть дополнен другими операциями, неперечисленными в дальнейшем. СИС 79 содержит процедуры, связанные со статистической обработкой данных. Процедуры сугубо математическо-статистического анализа не включены в СИС 79. Для решения таких проблем имеется много программ в распоряжении пользователей. /Благодаря гибким возможностям пользования системой Генера, эти процедуры без особых трудностей могут быть вызваны из программы./ При создании системы основной целью явилось обеспечение в легкообрабатываемой форме решения общих статистических задач и подготовительных операций математико-статистического анализа /контроля, кодирования, селекция, группирование, составление таблиц/. Надо отметить, что СИС 79 является не только языком для управления данными, но также обеспечивает эффективную обработку большой массы данных.

СИС 79 содержит задачи следующего типа:

1. Операции передачи и конверсии;
2. Топологическое отображение /контроль, кодирование/
и селекция;

3. Анализ статистических проб, создание основных данных для анализа математической статистики и составления таблиц;
4. Вывод форматов /таблиц/,
Интерфейсы для статистических оценок и для передачи данных в другие системы;

В дальнейшем подробно представляем эти задачи.

2.1. Операции передачи и конверсии

Программа СИС 79 служит прежде всего для обработки большой массы данных, поэтому эффективность операций ввода и вывода играет в ней большую роль.

СИС 79 содержит четыре операции ввода и вывода:

LECTOR

SCRIPT

SOBRED

SOBWRT

2.1.1. Команда LECTOR

Задачей команды является считывание ввода и конверсия на двоичную форму. С командой /в настоящем варианте/ производится считывание положительных целых чисел.

Расположение данных входной записи задается пользователем в форматах языка Фортран. Структура записи

определяется описанием, применяемым в КОБОЛ. Таким образом обеспечивается возможность совместного обращения с концентрированными данными. В отличие от КОБОЛ, здесь на каждом уровне могут задаваться форматы.

Целью команды является обеспечение быстрой конверсии. Это достигается эффективной процедурой, созданной для обработки ввода с форматами языка Фортран. В зависимости от структуры записи генерирования процедура ввода может даже на несколько порядков быстрее работать, чем традиционные процедуры в Фортране.

2.1.2. Команда SCRIPT

Задачей команды является печать данных, записанных в памяти, по форматам языка Фортран. Параметрическое изображение полностью совпадает с параметрическим изображением команды **LECTOR**. Первичной целью команды при этом является тоже ускорение конверсии.

2.1.3. Команда COBWRT

С использованием команды COBWRT данные могут быть переписаны на файл в уплотненной двоичной форме. Данные, записанные таким образом, могут быть считаны с файла с помощью команды COBRED.

При накоплении данных в формате символов потребность в памяти меньше чем при накоплении в двоичной форме /особенно в случае маленьких чисел/, но при этом исполнение расчета замедляется из-за конверсии. Процедура COBWRT обеспечивает запоминание неотрицательных целых чисел в уплотненной и в то же время в двоичной форме, необходимой для быстрой обработки. В каждом машинном слове помещается несколько данных. Каждый из данных занимает столько разрядов, сколько необходимо для запоминания его максимального значения.

При пользовании этой процедурой пользователь кроме названия печатаемых данных задает также ограничение по величине данного /или длину в разрядах/. На основе этого процедура распределит данные в уплотненной форме.

2.1.4. Команда COBRED

Командой производится считывание уплотненной двоичной записи, произведенной командой COBWRT. Пользователь задает описание такой же записи, как в команде COBWRT и на основе этого процедура производит первоначальные данные.

2.2. Операции топологического отображения и селекции

Эта группа операций содержит все задачи подготовки данных, требующие операции отображения /кодирование, анализ функций/. Такими являются простые и сложные задачи проверки данных, операции селекции и принятия решений. Эти задачи исполняются следующими командами:

BOUNDS

GRAPH

DECISE

При каждой операции отображения необходимо каким-то образом описать процедуру. Это не сложно, если речь идет о функции, заданной формулой, но в случае сложных функций, которые могут быть описаны только при помощи таблицы значений это может привести к таблицам громоздкого размера. Описание, проверка и запоминание таких таблиц очень сложная проблема. В системе СИС 79 эта проблема решается при помощи специфичной системы запоминания и подпрограммы "Мешок".

2.2.1. Подпрограмма "Мешок"

Функцией подпрограммы является реализация процедур для удобного заполнения таблиц /цифровой перечень слов, таблица переходов, таблицы значений функций/ в центральной памяти. Применение программы обеспечивает следующие выгоды:

1. Пользователь может описать таблицы значений в самой удобной, наглядной и плотной форме.
2. Описание таблиц значений синтаксически проверяется подпрограммой и дается подробный сигнал распознавания ошибок.
3. Обеспечивается быстрое и удобное отображение некоторых многопеременных функций.

2.2.2. Команда BOUNDS

С помощью этой команды проверяется, что значение переменных не выходит ли из заданных ограничений. Это может являться начальным шагом контроля данных и подготовкой к применению нижеописываемой команды **GRAPH**. Генерированная программа информирует пользователя о погрешностях в данных при помощи переменного счетчика ошибок и массива сбора кодов проверки.

2.2.3. Команда GRAPH

С помощью этой команды может быть создан набор программ для оценки таблиц перекодирования и проверки. Программа проверяет данные на основе таблицы, заполненной с помощью подпрограммы Мешок, а также находит ошибочные комбинации значений и определяет значение функции, относящейся к безошибочным комбинациям. С использованием этой команды заодно можно заполнить таблицу, так как она непосредственно связана с подпрограммой Мешок. При прогоне генерированной программы /также как в случае команды BOUNDS/ пользователь получит информацию о данных с погрешностями при помощи счетчика ошибок и массива сбора кодов проверки.

Таблица проверки и перекодирования служит для описания целочисленных функций и применяется значительно эффективнее, чем традиционные таблицы значений, если в таблице значений некоторые части повторяются.

Функция описывается с помощью иерархического графа, размещенного на уровнях по отдельным переменным: на первом уровне решение принимается в зависимости от значения первой переменной, на втором уровне - от значения второй переменной и т.д. Каждый

уровень включает в себя один или несколько частичных таблиц /мешков/, содержащих информацию о той части возможных значений переменной соответствующего уровня, которая попадет в один интервал. При этом получена информация о том, что функция существует ли на данное значение, какого её значение, а также получена команда по дальнейшему анализу.

2.2.4. Команда DECISE

При статистической обработке вместо интегрирования системы данных часто целесообразно прибегать к разделению на подсистемы по различным информационным потребностям. Аспекты деления часто описываются с весьма сложными логическими выражениями. Команда DECISE облегчает описание сложных логических выражений и осуществляет эффективную оценку логических выражений. Логические функции, описанные таким образом могут применяться не только при селекции статистических множеств данных, но также при других задачах принятия решений.

2.3. Анализ статистических проб

При статистическом анализе требуются данные не только об индивидуумах, но также о характере групп. Взятие пробы производится по индивидуумам, поэтому при статистическом анализе одним из самых важных и характерных шагов является группировка и суммирование данных, определение частоты случаев. Это служит основой для общего анализа и для математической статистики.

В СИС 79 могут быть собраны частоты, а также сумма значений одной и более выбранных переменных /например, суммы произведений и квадратов для составления корреляционной матрицы/.

В системе обеспечивается обработка большой массы данных при помощи быстрой процедуры сбора.

2.4. Составление таблиц

В СИС 79 составление статистических таблиц производится на основе заранее собранных частот и сумм, поэтому для составления каждой таблицы /в зависимости от размера таблицы/ требуется только несколько секунд машинного времени.

SUMMARY

SYSTEM SIS79/GENERA

P. Kerékfy

Generating programs is a well-known method for easy and efficient programming. Our system SIS79 is based on generator system GENERA. GENERA integrates any number of generator subsystems into a single precompiler. A file of host language statements and directives is processed to create the host language program.

Statistical information system SIS79 contains a number of data handling and data modification tasks. Interfaces to statistical packages can be established.

A SIS79/GENERA rendszer

Kerékfy Pál

A programok generálása egy jól ismert módszer a programozásban. A SIS79 rendszert a GENERA generáló rendszer segítségével építettük fel. A GENERA tetszőleges számú generáló eljárást egyesít egyetlen előfordítóvá. A rendszer alapnyelven írt forrásprogramban elhelyezett utasítások alapján hozza létre a kívánt programot.

A SIS79 statisztikai információs rendszer adatkezelő, konverziós és függvényeket kiértékelő eljárásokat tartalmaz. A különböző statisztikai programcsomagokhoz könnyen kapcsolható.