

Ц 840
А-85

Н 15
Центральная библиотека

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ
ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ И АВТОМАТИЗАЦИИ

71/3
На правах рукописи

И.С. Лупашина

ТРАНСЛЯЦИЯ С ФОРТРАНА НА МАШИНАХ
С МАЛЫМ ОБЪЕМОМ ОПЕРАТИВНОЙ ПАМЯТИ

А в т о р е ф е р а т
диссертации, представленной на соискание ученой степени
кандидата физико-математических наук

Дубна 1971

Работа выполнена в Институте физики высоких энергий,
г. Серпухов.

Научный руководитель - кандидат физико-математических
наук С.Н.Соколов.

Официальные оппоненты: доктор физико-математических
наук Н.Н. Говорун, кандидат физико-математических наук
А.Н.Томилин.

Ведущее предприятие - ВЦ АН СССР.

Автореферат разослан "12" января 1971 г.

Защита диссертации состоится "12" марта 1971 г.
на заседании Ученого Совета ОИЯИ.

С диссертацией можно ознакомиться в библиотеке ОИЯИ.

Ученый секретарь Совета



Ю.Катьшев

И.С. Лупашина

4539 в.р.
6851

ТРАНСЛЯЦИЯ С ФОРТРАНА НА МАШИНАХ
С МАЛЫМ ОБЪЕМОМ ОПЕРАТИВНОЙ ПАМЯТИ

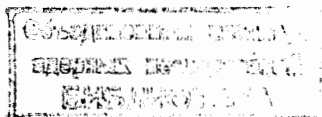
010 – математическое программирование

А в т о р е ф е р а т

диссертации, представленной на соискание ученой степени
кандидата физико-математических наук

Научный руководитель –
кандидат физико-математических наук

С.Н.Соколов



В октябре 1967 года в Институте физики высоких энергий (г. Серпухов) сдана в эксплуатацию Мониторная система ИФВЭ-67 - система автоматизации программирования на основе языка ФОРТРАН^{/5-7,11,12/} для машины Минск-22. К апрелю 1970 года были закончены работы по написанию и отладке Мониторной системы ИФВЭ-68 для ЭВМ М-220^{/1,2,12/}.

В Систему входят диспетчерские, загрузочные, контрольно-диагностические программы, программы ввода-вывода, библиотека стандартных программ и транслятор. Собственно транслятор составляет несколько менее половины всей Системы.

Выбор ФОРТРАНа в качестве основного языка для автоматизации программирования в институте был обусловлен следующими причинами. Во-первых, за рубежом существует большая библиотека написанных на ФОРТРАНе программ, значительно превосходящая по объему библиотеки программ на других алгоритмических языках. Трудоемкость создания только программ обработки экспериментальных данных, являющихся наиболее важными для института, во много раз превосходит трудоемкость создания транслятора с ФОРТРАНа. Во-вторых, в языке ФОРТРАН заложена возможность работы с большими комплексами относительно самостоятель-

ных программ (сепаратная трансляция подпрограмм, наличие оператора COMMON и т.п.). В-третьих, ФОРТРАН содержит развитую группу операторов ввода-вывода по форматам, прост для усвоения и относительно легко допускает вставки в программы блоков, написанных на языках, близких к машинному коду. Таким образом, язык ФОРТРАН может служить единой основой при программировании практически любых задач, возникающих в институте. (Эти же соображения заставили сотрудников ОИЯИ (г. Дубна) писать транслятор с ФОРТРАНа для ЭВМ БЭСМ-6^{13,14}).

Для сравнения заметим, что близкий к ФОРТРАНУ язык АЛГОЛ в его настоящем виде¹⁵ не может служить основой для программирования больших комплексов программ, так как по перечисленным выше пунктам он уступает ФОРТРАНУ.

В данной работе дается анализ проблем, возникающих при написании транслятора с ФОРТРАНа для машин среднего класса с малым объемом оперативной памяти (8К). Для простоты и конкретности изложения в основной части работы речь идет лишь о ТФ для ЭВМ М-220.

1. Транслятор как часть Мониторной системы

Транслятор является специализированной секцией Мониторной системы, осуществляющей перевод отдельной (под)программы (функции) с входного языка на язык сборки (Бибкод) – компактную библиотечную форму, которая кроме машинных команд в относительных адресах содержит дополнительную информацию для Мониторной системы. Объединение (под)программ (функций), транслированных независимо друг от друга, в единую рабочую программу, полностью готовую к счету, не является функцией транслятора и осуществляется другой секцией Мониторной системы.

Вошедший в состав Мониторной системы ИФЗ-68 для ЭВМ М-220^{1,2} транслятор реализует наиболее полный вариант языка

ФОРТРАН - расширенный АСА ФОРТРАН^{/3,8/}. К последнему, в соответствии с языком ЦЕРН ФОРТРАН^{/4/}, добавлены операторы PRINT, PUNCH, расширены операторы READ, WRITE, введены маскирующие выражения, добавлен новый оператор LOAD, управляющий загрузкой секций в оперативную память^{/5-7/}. Из операторов АСА ФОРТРАНа не реализован оператор BLOCK DATA.

Для максимального использования возможностей машины наряду с ФОРТРАНОм в произвольной смеси с ним в качестве входного языка допускается Полукод - язык из семейства Автокодов 1 : 1. В связи с тем, что Полукод является еще и промежуточным языком при трансляции, множество элементов, на котором он определен, включает в себя как понятия языка ФОРТРАН (идентификаторы, метки), так и понятия, генерируемые в процессе трансляции (ячейки RETURN, STOP, END, абсолютные адреса (nA), информационные ячейки (nI), рабочие ячейки (nR), ячейки типа n+, n-, названия констант (nC), искусственные метки (nB, nE, nF)). Основной единицей Полукода является строка Полукода, состоящая из пяти шестисимвольных слов. Признаком, определяющим вид оператора Полукода, служит комментарий (второй символ второго слова). В зависимости от вида комментария (M, I, V, H, P, C, L, X, R, D, S, E, T, A) по-разному будет понята информация, содержащаяся во втором, третьем, четвертом и пятом словах оператора.

Оттранслированная (под)программа (функция) имеет вид стандартного массива (Библиотечного кода). Часть информации, входящей в состав Бибкода, необходима Системе во время сборки (таблица ENTRY-входов, таблица названий и длинCOMMON-блоков, тело программы), часть же служит лишь для целей диагностики (на уровне входного языка) в процессе счета по собранной программе.

2. Структура транслятора

ТФ для ЭВМ М-220 принадлежит к разряду классических с некоторыми элементами, присущими синтаксически ориентированным трансляторам /9/.

Трансляция осуществляется по локальному принципу. Это означает, что каждому выполняемому оператору на языке ФОРТРАН после трансляции соответствует группа машинных команд, и последовательность операторов совпадает с последовательностью этих групп. Исключение составляет оператор DO, который в результате трансляции превращается в группу команд в начале и одну команду в конце цикла. Экономия команд производится локальным образом в пределах нескольких команд, относящихся к одному оператору.

Трансляция (под)программы (функции) распадается на две фазы: перевод с входного языка в Полукод и перевод с Полукода в Бибкод.

Во время первой фазы трансляции происходит следующее. С магнитного барабана в оперативную память считывается максимально возможная порция программы на входном языке. Отдельный оператор (если он фортрановский) распознается с помощью специальной программы-анализатора. Распознанный оператор передается соответствующей специализированной программе, которая и транслирует его в Полукод. (Под) программа (функция) в Полукоде, получившаяся после первой фазы трансляции, включает в себя всю необходимую для построения рабочей программы информацию, которую содержала (под) программа (функция) на языке ФОРТРАН. Она может быть выведена на АЦПУ, может быть отперфорирована и использована вместо исходной (под)программы (функции) в качестве материала для трансляции.

Однако, обладая свойством полноты и наглядности, Полукод не приспособлен для многократных просмотров. Сведения, содержащиеся в нем, "размазаны" по всей программе и зачастую дублируют друг друга. Поэто-

му информацию, влияющую на ход трансляции, лучше всего хранить табличным способом, а Полукод рассматривать как источник для заполнения таких таблиц, последние имеют списочную структуру^{/10/}.

В процессе перевода программы с ФОРТРАНа в Полукод в целях реализации отдельных операторов генерируются новые понятия. Для построения этих понятий также необходимы некоторые таблицы, необходимость в которых отпадает по окончании этой фазы перевода (таблицы для построения искусственных меток, таблицы констант, таблицы рабочих ячеек).

Во время второй фазы трансляции полукодовая программа порциями считывается в оперативную память, выполняемые полукодовые строки заменяются машинными командами в относительных адресах, дополнительная информация для Мониторной системы выделяется из таблиц и заносится в стандартный массив. Если во время трансляции не было обнаружено ошибок, стандартный массив записывается на отладочную магнитную ленту DEBUGG рядом с программой на входном языке.

Информация об ошибках сообщается программисту. Она содержит номер ошибки, номер строки входного текста, вид ошибки, текст оператора, содержащего ошибку.

3. Проблема распознавания операторов

Распознавание оператора производится по принципу сверки текста оператора с неким набором правил (структур). Каждому оператору соответствует одна или несколько таких структур (когда оператор имеет несколько допустимых вариантов записи). Подходящую структуру подпрограмма STATEMENT ищет, пробуя их в том порядке, в каком они помещены в таблице. Оператор замещения намеренно написан последним, так как его трудно отвергнуть по первым нескольким символам. Параллельное движение по тексту оператора и структуре организовано по принципу:

раньше пробуются тот вариант соответствия, который позволяет продвинуться дальше по тексту оператора, меньше продвинувшись по тексту структуры.

Такой подход к проблеме распознавания позволяет на самой ранней стадии трансляции с большой степенью точности выявить формальные (с точки зрения грамматики) ошибки в написании отдельных операторов и выдать информацию о них. Тогда на последующих стадиях останется только выявить ошибки, найти которые возможно, лишь зная взаимосвязь отдельных операторов.

В языке ФОРТРАН легко отделить операторы (или части операторов), которые просто формализуются (так как имеют довольно жесткий синтаксис), от операторов, формализовать которые значительно труднее (из-за большого разнообразия допустимых возможностей). Полный набор правил, отвечающий таким операторам, велик по объему и достаточно сложен, он занимает значительное место в памяти, что крайне нежелательно при ограниченном объеме последней. Поэтому подробность структур, отвечающих различным операторам, может быть выбрана различной. Для легко формализуемых операторов это может быть полный набор правил, для остальных - в правилах должны быть отражены лишь характерные черты, которые достаточны для опознания оператора, а подробный контроль структуры таких операторов вынесен в подпрограммы, фактически осуществляющие их трансляцию.

В качестве иллюстрации записи правил грамматики приведены два примера. В первом рассмотрено написание структуры для легко формализуемого оператора (одного из самых сложных среди операторов этого класса), во втором - написание точной структуры для трудно формализуемого оператора, и параллельно приведено приемлемое компромиссное решение для этого оператора.

Структура состоит из отдельных элементов, каждый из которых в наглядной записи имеет вид:

a
b
c

где a - код символа или номер одной из следующих таблиц: l - букв; n - десятичных цифр; p8 - восьмеричных цифр; ln - букв и десятичных цифр; e - букв, десятичных цифр и разделителей + - * (.,); io - букв, десятичных цифр и разделителей + - * (, =); f - букв, десятичных цифр и разделителей + - * / . , (=); b - либо (- открывающая структурная скобка объединения элементов,) - закрывающая структурная скобка, либо признак, указывающий на то, что "a" является: s - конкретным символом, t - символом из таблицы; c - режим присутствия данного элемента (или объединения элементов, заключенных в структурные скобки), означающий: ≥ 0 - элемент (объединение) может либо не присутствовать, либо присутствовать любое число раз; ≤ 1 - элемент (объединение) может присутствовать один раз или не присутствовать вообще; = 1 - элемент (объединение) обязательно присутствует и только один раз; ≥ 1 - элемент (объединение) обязательно присутствует не менее одного раза; r - (рекурсия), если объединение, содержащее элемент (объединение), помеченный режимом r, присутствует несколько раз, то в первый раз этот элемент (объединение) отсутствует, а в последующие обязательно присутствует; / - (взаимоисключающее "или") ставится только у открывающей структурной скобки и означает, что выражение в скобках есть список, состоящий из элементов или их объединений, один и только один член которого обязательно присутствует.

Данный метод записи структур по своей идее ближе всего к нерекурсивной форме синтаксических обозначений, описанной в ^{16-18/}. Он оказался практически удобным для построения компактной и относительно быстрой программы синтаксического контроля.

З а к л ю ч е н и е

При создании трансляторов с ФОРТРАНа для ЭВМ Минск-22 и М-220 в основу был положен метод раскручивания^{/9/}, сущность которого заключается в следующем. Предположим, что для некоторой машины Р существует транслятор с проблемно-ориентированного языка L, в терминах которого может быть определен процесс трансляции. Требуется создать транслятор с языка М для машины R. Выполнение следующих шагов обеспечит решение данной задачи:

1. Записать на языке L программу, определяющую процесс трансляции с языка М в код машины R ;
2. транслировать эту программу, используя машину Р ; получим транслятор с языка М , который работает на машине Р , но выдает результат трансляции в коде машины R ;
3. записать на языке М программу, определяющую процесс трансляции с языка М в код машины R ;
4. транслировать эту программу на машине Р , используя транслятор, полученный на шаге два; в результате получим транслятор с языка М , который работает на машине R .

В структуре ТФ обращает на себя внимание тот факт, что из транслятора легко выделяется часть, являющаяся ПП с Полукода. Транслятор с Полукода и часть Мониторной системы, ведающая хранением, корректурой и сборкой программ, были составлены и отлажены в машинном коде Минска-22. Это позволило производить всю дальнейшую работу методом раскручивания. При создании ТФ для Минска-22 в качестве машин Р и R выступала ЭВМ Минск-22, в качестве языка L - Полукод, а языка М - ФОРТРАН ИФВЭ. При создании ТФ для М-220 машинами Р и R служили Минск-22 и М-220, соответственно, в качестве языка L выступал ФОРТРАН ИФВЭ, а языка М - расширенный АСА ФОРТРАН.

Использование метода раскручивания при создании ТФ для ЭВМ М-220 было осложнено тем, что машины Минск -22 и М-220 принадлежат к различным типам (различная длина ячейки, адресность, представление констант и т.д). Это привело к тому, что транслятор, написанный на шаге один (ТАНДЕМ), оказался более сложным, чем автономный транслятор для М-220 (так как на первом шаге приходилось учитывать особенности обеих машин).

При написании трансляторов с ФОРТРАНа не было цели достичь максимального быстродействия всех программ. Основная работа по оптимизации была проведена во время отладки, при этом систематически использовался следующий прием. Во время трансляции машину многократно останавливали в случайные моменты времени и отмечали подпрограммы транслятора, работавшие в эти моменты. Накопленная статистика выявила несколько подпрограмм, расходующих почти все время трансляции. В эти подпрограммы вносились изменения, значительно (обычно в несколько раз) ускоряющие работу. Оптимизация, выполненная таким образом, дала весьма ощутимый эффект при очень небольшом объеме переделок (несколько процентов текста).

Транслятор с ФОРТРАНа на Минске-22 интенсивно эксплуатируется с сентября 1967 года. Использование полученных из-за рубежа программ показало, что отсутствие в языке ФОРТРАН ИФВЭ некоторых операторов не вызывает особых затруднений, так как эти операторы легко заменяются эквивалентами, составленными из имеющихся.

Более серьезным препятствием является малый объем оперативной памяти машины и, как следствие, острая нехватка места для таблиц. Составленные в расчете на большие машины, подпрограммы часто содержат слишком много блоков COMMON, массивов, меток и переменных. Сокращение числа массивов и разбиение подпрограмм на более мелкие требует значительного труда и вносит ошибки. Поэтому при разработке транслятора для ЭВМ М-220 кроме расширения входного языка значительное внимание было уделено увеличению емкости таблиц.

Основные характеристики трансляторов с ФОРТРАНа для ЭВМ Минск-22 и М-220 приведены в таблице.

Т а б л и ц а

ТФ для ЭВМ Минск-22	ТФ для ЭВМ М-220
<p>Транслятор входит в состав Мониторной системы ИФВЭ -67. Работа по созданию этой Системы была начата в апреле 1966 года и велась группой из шести человек. В 1966 г. были составлены на ФОРТРАНе рабочие тексты транслятора и других секций Мониторной системы, переписаны в код машины, отлажены черновой вариант Системы и транслятор с промежуточного языка Полукода. Дальнейшие работы по расширению Мониторной системы, отладке полного транслятора с языка ФОРТРАН ИФВЭ и минимальной библиотеки программ производились с помощью услуг самой Системы, возможности которой постепенно возрастали вплоть до окончания работ в сентябре 1967 года.</p>	<p>Транслятор входит в состав Мониторной системы ИФВЭ-68. Работа по созданию этой Системы была начата в апреле 1968 года и велась группой из шести человек. К апрелю 1969 года был составлен на ФОРТРАНе и отлажен ТАНДЕМ (вариант Системы, отвечающий первому шагу раскрутки). К апрелю 1970г года были закончены работы по написанию и отладке автономного варианта Мониторной системы ИФВЭ-68 для ЭВМ М-220.</p>
<p>Транслятор состоит из 100 подпрограмм, написанных на ФОРТРАНе со вставками Полукода. Подпрограммы транслятора, зависящие от системы команд машины, насколько это возможно, отделены от остальных. Тексты программ транслятора не содержат одинаковых или близких блоков: в процессе написания они выделялись в отдельные подпрограммы.</p>	

Транслятор с помощью оператора LOAD разбит на 5 секций. В первую секцию включены программы, общие для первой и второй фаз трансляции, во вторую - программы, осуществляющие трансляцию из Полукода в стандартный массив; в третью секцию - программы, общие для трансляции всех операторов, в четвертую - программы, транслирующие в Полукод заголовочные операторы и операторы описания; в пятую секцию включены программы, осуществляющие трансляцию в Полукод всех остальных операторов языка ФОРТРАН ИФВЭ.

Транслятор с помощью оператора LOAD разбит на 15 секций. В первую секцию включены программы, общие для первой и второй фаз трансляции; в секции со второй по пятую - программы, осуществляющие трансляцию из Полукода в стандартный массив; в шестую секцию программы, общие для трансляции всех операторов; в седьмую - программы, переносящие константы из таблицы TYPE в Полукод; в секции с восьмой по четырнадцатую - программы, транслирующие в Полукод все операторы ФОРТРАНа; в пятнадцатую секцию включены программы, осуществляющие распознавание операторов.

Общая длина транслятора 12000 ячеек, из них таблицы и рабочие поля занимают 4000 ячеек. Средняя скорость трансляции 40 команд в минуту.

Трудоемкость работы по написанию и отладке транслятора составляет 4,5 человеко-года, не считая труда инженеров и лаборантов.

Общая длина транслятора 18000 ячеек, из них таблицы и рабочие поля занимают 5500 ячеек. Средняя скорость трансляции 90 команд в минуту.

Трудоемкость работы по написанию и отладке практически двух трансляторов (ТАНДЕМа и автономного) составляет 3 человеко-года, не считая труда инженеров и лаборантов.

Работы по дальнейшему усовершенствованию трансляторов с ФОРТ-РАНа для машин с малым объемом оперативной памяти будут направлены на улучшение их эксплуатационных характеристик:

1. увеличение скорости трансляции (главным образом, за счет отказа от принципа полноты и установления более естественной связи между Полукодом и таблицами);

2. улучшение качества трансляции (за счет улучшения трансляции индексных выражений и др.).

Реализация Мониторной системы на машинах класса Минск-22 и М-220 подтвердили осуществимость и целесообразность принятой в ИФВЭ программы обеспечения всех вычислительных машин ИФВЭ автоматизированными системами программирования на основе языка ФОРТРАН.

В заключение автор выражает глубокую благодарность С.Н.Соколову за научное руководство в работе. Автор признателен Л.В.Беляевской, П.А.Калинченко, А.С.Маркову, В.А.Макарову, И.В.Поповой, А.С.Лузиной - соавторам, Л.Д.Харыбиной, оказавшей неоценимую помощь в отладке, Л.Рыжовой за техническую помощь в работе.

Л и т е р а т у р а

1. П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова, С.Н.Соколов. Препринт ИФВЭ 68-41-К, Серпухов, 1968.
2. В.Д.Жильченков, А.С.Марков, В.Д.Матвеев, С.Н.Соколов. Препринт ИФВЭ 69-33, Серпухов, 1969.
3. FORTRAN vs. Basic FORTRAN Comm. ACM, 7, 10, 591-625, 1964.
4. CERN-6600. Computer. CERN Fortran. Издание ЦЕРН, Женева, 1964.
5. С.Н.Соколов, П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова. Препринт ИФВЭ 67-41-К, Серпухов, 1967.
6. С.Н.Соколов, П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова. Препринт ИФВЭ 68-30-К, Серпухов, 1968.
7. П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова, С.Н.Соколов. АН СССР. Журнал вычислительной математики и математической физики, т. 10, № 2, М., 1970.
8. Appendices to ASA FORTRAN. Comm. ACM., 8, 5, 287-288, 1965.

9. П.Ингерман. Синтаксически ориентированный транслятор. М., 1969.
10. Г.И.Бабецкий, М.М.Межанова, Ю.М.Волошин, А.П.Ершов и др. АЛЬФА-система автоматизации программирования. ВЦ СО АН СССР, 1967.
11. С.Н.Соколов, П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова. Транслятор с языка ФОРТРАН ИФВЭ. Доклад на Первой Всесоюзной конференции по программированию. Киев, ноябрь, 1968.
12. С.Н.Соколов, Л.В.Беляевская, П.А.Калинченко, И.С.Лупашина, В.А.Макаров, А.С.Марков, И.В.Попова. ФОРТРАН и Мониторная система. Изд-во "Статистика", М., 1970.
13. З. Бродцински, В.Ю.Веретенев, Гизе Петер, Гизе Пирושка, Р. Гирр, Н.Н.Говорун и др. Транслятор с языка Фортран для системы математического обеспечения БЭСМ-6. Доклад на Первой Всесоюзной конференции по программированию. Киев, ноябрь, 1968.
14. В.Ю.Веретенев, Н.Н.Говорун, Е.А.Жоголев и др. Вариант операционной системы для серийного образца машины БЭСМ-6. Доклад на Первой Всесоюзной конференции по программированию. Киев, ноябрь, 1968.
15. С.С.Лавров. Универсальный язык программирования (АЛГОЛ-60). Изд-во "Наука", М., 1964.
16. K.E.Iverson. A method of syntax specification. Comm. ACM, 1964, 7, N10, 588-589.
17. W.H.Burkhardt. Metalanguage and syntax specification. Comm. ASA, 1965, 8, N5, 304-305.
18. J.W.Carr, J.Weiland. A nonrecursive method of syntax specification. Comm. ACM, 1966, 9, № 4, 267-269.

Рукопись поступила в издательскую группу
24 декабря 1970 года.