

Б-833

На правах рукописи

Бородич Лилия Ивановна

ВОПРОСЫ КОНСТРУИРОВАНИЯ СИСТЕМ
ОТЛАДКИ (ПРАГМАТИЧЕСКИЙ ПОДХОД)

Специальность 01.01.10 – математическое обеспече-
ние вычислительных машин
и систем

А в т о р е ф е р а т

диссертации на соискание ученой степени
кандидата физико-математических наук

ДУБНА – 1978

Работа выполнена в лаборатории автоматизации программирования Института математики АН БССР.

Научные руководители -

доктор физико-математических наук,
профессор В.Н.РЕДЬКО

кандидат физико-математических наук,
старший научный сотрудник - Н.В.ШКУТ

Официальные оппоненты -

член-корреспондент АН УССР, доктор физико-математических наук, профессор В.Л.КЩЕНКО
(Институт кибернетики АН УССР, г.Киев)

кандидат технических наук, старший
научный сотрудник Г.В.ПЕЛЕДОВ
(НИЦЭВТ, г.Москва)

Ведущее предприятие -

Научно-исследовательский вычислительный
центр МГУ им. М.В.Ломоносова (г.Москва)

Защита диссертации состоится 12 мая 1978г.
в 11 часов на заседании специализированного совета Д047.01.04
при Лаборатории вычислительной техники и автоматизации Объединенного Института ядерных исследований, г.Дубна, Московской области.

С диссертацией можно ознакомиться в научно-технической библиотеке ОИЯИ.

Автореферат разослан "4" апреля 1978 года

Ученый секретарь специализированного
совета, кандидат физ.-мат. наук

Т.П.ПУЗЫНИНА

Пузынина

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность. Как известно, 60-70% затрат вычислительных центров составляют затраты на создание программ (программирование, компиляция, отладка). Существенная часть этих затрат идет на отладку программ, что определяет эффективность программирования и решения задач на ЭВМ.

Значительный удельный вес составляет первый этап отладки - синтаксический и семантический контроль программы в процессе компиляции, разработке методов которого и их реализации посвящена диссертационная работа.

Цель работы. Разработка и реализация системы контроля программ в языках типа АЛГОЛ в процессе компиляции.

Научная новизна и практическая ценность. Предложен и реализован метод распознавателей, обеспечивающий решение вопросов локализации и нейтрализации ошибок в условиях оптимизирующего транслятора серийных машин.

Реализация результатов. Предложенный метод реализован в трансляторах ТАМ-2(22) для ЭВМ "Минск-2(22)", ТАМ-22Т для ЭВМ "Минск-22М", ТАМ-32 для ЭВМ "Минск-32", АЛГАМС ДОС для ЕС ЭВМ, широко используемых в практике решения задач.

Апробация работы. Результаты диссертации докладывались на ряде конференций и семинаров: конференции математиков Белоруссии (1968, 1971, 1975гг.), Всесоюзная межвузовская конференция по математическому обеспечению автоматизированных систем управления (1970г.), научно-техническая конференция "Проблемы создания и совершенствования технических и программных средств ЕС ЭВМ" (1977г.), Всесоюзные семинары Ассоциации пользователей машин типа "Минск", семинар Киевского государственного университета, семинар ЛВТА г.Дубна, семинар ВЦ МГУ, семинары ИМ АН БССР.

Публикации. По материалам диссертации опубликовано 17 работ (в том числе одно учебное пособие и два руководства (в соавторстве)).

Объем работы. Диссертация состоит из введения, трех глав, заключения и содержит 148 страниц машинописного текста, включая 10 рисунков и 4 таблицы в тексте, список литературы, насчитывающий 83 наименования и два приложения.

Общедоступный документ
Института математики
Академии наук БССР

Введение посвящено краткому изложению постановки задачи и обзору результатов.

Первая глава "Характеристика методов разбора и их оценка с точки зрения практической реализации" посвящена обзору и сравнительному анализу методов синтаксического разбора, определение на основе этого анализа постановки задачи, состоящей в разработке эффективного метода синтаксического анализа для целей контроля (названного в работе методом распознавателей). Отличительная особенность метода составляет разделение процесса распознавания согласно структуре синтаксического дерева на ряд слоёв, каждому из которых соответствует совокупность распознавателей некоторых определяющих метапеременных. Такое расслоение алгоритма анализа обеспечивает высокую скорость в его выполнении за счет уменьшения числа обращений. Основу метода составляет выделение распознавателей, общих для различных ветвей синтаксического дерева, и их реализация в виде процедур (в общем случае рекурсивных), в чем выражается рекурсивная структура языка.

Вторая глава "Мотивировка принятого метода контроля с точки зрения условий конкретных реализаций" посвящена выработке и обоснованию требований к методам синтаксического контроля и семантической правильности программ применительно к конкретным реализациям. С этой целью определяются требования к программному обеспечению ЭВМ широкого применения и, в частности, требование простоты использования, наличия средств диагностики.

Выясняются особенности языков типа АЛГОЛ, определяющие структуру анализаторов (наличие рекурсии, блочной структуры, областей действия описаний идентификаторов, множественность порождающих правил). Кроме того, формализуются некоторые семантические свойства (не охватываемые Бэкусово-Науровскими формами) такие как: единственность описаний идентификаторов в блоке, обязательность описания идентификаторов, связь идентификаторов с областью их действия, совпадение числа фактических и формальных параметров процедур, соответствие классов величин и согласование типов операндов, использование идентификаторов в соответствии с их описанием. Обосновывается выбор аппарата бинарных отношений для реализации общего подхода к синтаксическому и семантическому

контролю программ. Мотивируется выделение системы отладки в автономный блок синтаксического и частичного семантического контроля, избавляющий собственно компилятор от функций контроля и компиляции заведомо некорректных программ.

В этой же главе рассмотрены особенности реализации метода распознавателей в трансляторах ТАМ-2(22), ТАМ-22Т, ТАМ-32, АЛГАМС ДОС ЕС ЭВМ. В трансляторах ТАМ-2(22) реализована общая идея метода распознавателей; в ТАМ-22Т осуществлена минимизация размера текста, пропускаемого при обнаружении единичной ошибки, и полная локализация единичных ошибок, а также определение их характера (типа) и вида (степени грубости); системы отладки трансляторов ТАМ-32, АЛГАМС ДОС ЕС, кроме того, реализованы строго по модульному принципу.

Третья глава "Метод распознавателей и его реализация" посвящена детализации и реализации предложенного метода построения системы отладки. Согласно этого метода за один просмотр автономно от собственно компиляции осуществляется контроль исходной программы по смешанной стратегии - безвозвратной нисходящей с элементами восходящей стратегии. При этом осуществляется полный лексический и синтаксический контроль и выявляется существенный класс наиболее характерных семантических ошибок.

Одну из ведущих идей реализации метода составляет формализация некоторых свойств посредством аппарата бинарных отношений, задаваемых в виде таблиц стандартного типа, что позволяет реализовать блок отладки в виде метаалгоритма, не привязанного жестко к конкретным свойствам языка. К указанным таблицам относятся:

- таблица меток, потребность в которой вызвана возможностью использования метки до ее описания, запретом переходов извне в блоки и циклы и на неопределенные метки;
- таблица фактических и формальных параметров для проверки совпадения их количества при наличии фактических параметров, включающих суперпозиции функций;
- таблица размерностей массивов для проверки соответствия количества индексов, задаваемых выражениями, содержащие индексные выражения;
- таблица блочной структуры для контроля баланса открывающих и закрывающих операторных скобок и управления локализацией идентификаторов;

- таблица процедур-функций для контроля наличия присваивания идентификатору процедуры-функции.

Таблицы содержат характеристическую информацию, извлекаемую из описаний и спецификаций: класс, тип, размерность (для массивов), число параметров (для процедур), отношение к списку значений для формальных параметров и др. Все таблицы ведутся по строковому принципу. В работе установлена функциональная зависимость размеров таблиц от структуры исходной программы, позволяющие прогнозировать число обращений к внешним устройствам, минимизировать память под каждую из этих таблиц с целью ускорения анализа программ.

При построении анализаторов с целью уменьшения числа распознавателей (и тем самым повышения быстродействия алгоритма анализа) грамматике анализируемого языка может быть преобразована к виду, в котором число метапеременных сокращено за счет устранения некоторых нерекурсивно определяемых понятий, а также могут быть введены некоторые новые обобщающие метапеременные, например, общее выражение в языке АЛГАМС (см. таблицу № I).

Сокращение метапеременных иллюстрируется на примере определения программы.

1. $\langle \text{Программа} \rangle ::= \langle \text{Блок} \rangle | \langle \text{Составной оператор} \rangle$
2. $\langle \text{Блок} \rangle ::= \langle \text{Непомеченный блок} \rangle | \langle \text{Метка} \rangle : \langle \text{Блок} \rangle$
3. $\langle \text{Составной оператор} \rangle ::= \langle \text{Непомеченный составной} \rangle | \langle \text{Метка} \rangle : \langle \text{Составной оператор} \rangle$
4. $\langle \text{Непомеченный блок} \rangle ::= \langle \text{Начало блока} \rangle ; \langle \text{Конец составного} \rangle$
5. $\langle \text{Начало блока} \rangle ::= \textit{begin} \langle \text{Описание} \rangle | \langle \text{Начало блока} \rangle ; \langle \text{Описание} \rangle$
6. $\langle \text{Непомеченный составной} \rangle ::= \textit{begin} \langle \text{Конец составного} \rangle$
7. $\langle \text{Конец составного} \rangle ::= \langle \text{Оператор} \rangle \textit{end} | \langle \text{Оператор} \rangle ; \langle \text{Конец составного} \rangle$

Для проведения контроля программы потребовалось бы 7 распознавателей, исключая определение метапеременных $\langle \text{Оператор} \rangle$ и $\langle \text{Описание} \rangle$. Если же исходить из определяющих ее формул

1. $\langle \text{Программа} \rangle ::= \langle \text{Блок} \rangle | \langle \text{Составной оператор} \rangle$
2. $\langle \text{Блок} \rangle ::= \langle \text{Метка} \rangle \dots \textit{begin} \langle \text{Описание} \rangle ; \dots ; \langle \text{Описание} \rangle ; \langle \text{Оператор} \rangle ; \dots ; \langle \text{Оператор} \rangle \textit{end}$
3. $\langle \text{Составной оператор} \rangle ::= \langle \text{Метка} \rangle \dots \textit{begin} \langle \text{Оператор} \rangle ; \dots ; \langle \text{Оператор} \rangle \textit{end}$,

то потребуется 3 распознавателя.

Для языка АЛГАМС, например, были определены 29 распознавателей, перечень которых дан во втором столбце таблицы № I, а в её третьем столбце приведен список соответственно вызываемых ими распознавателей.

Таблица № I

№№ п/п	Имя вызывающего распознавателя	Имена вызываемых распознавателей
1	2	3
1.	Программа	Блок, оператор
2.	Блок	Описание, оператор, контролер ошибки
3.	Описание	Описания: типа, массива, процедуры, переключателя; контролер ошибки
4.	Оператор	Операторы: присваивания, условный, цикла, перехода, составной, блок, ввода, вывода, разметки; контролер ошибки
5.	Оператор присваивания	Выражения: арифметическое, булевское, индексное; контролер ошибки
6.	Условный оператор	Оператор, безусловный оператор, контролер ошибки
7.	Оператор перехода	Именуемое выражение, контролер ошибки
8.	Оператор цикла	Выражения: арифметическое, булевское; оператор, контролер ошибки
9.	Составной оператор	Оператор, контролер ошибки
10.	Описание типа	-
11.	Описание массива	Арифметическое выражение
12.	Описание процедуры	Спецификатор, оператор, ограничитель параметра, контролер ошибки
13.	Описание переключателя	-
14.	Арифметическое выражение	Выражения: арифметическое, булевское, индексное; функция, стандартная функция, контролер ошибки
15.	Индексное выражение	Арифметическое выражение, контролер ошибки
16.	Булевское выражение	Выражения: общее, булевское, индексное; функция, контролер ошибки
17.	Общее выражение	Выражения: булевское, арифметическое, общее, отношения; функция, контролер ошибки

1	2	3
18.	Безусловный оператор	Оператор, контролер ошибки
19.	Оператор процедуры	Список параметров, общее выражение
20.	Список параметров	Общее выражение, ограничитель параметра
21.	Функция	Список параметров, общее выражение, ограничитель параметра
22.	Выражение отношения	Выражения: арифметическое, булевское
23.	Именуемое выражение	Арифметическое выражение, контролер ошибки
24.	Оператор ввода	Выражения: арифметическое, индексное
25.	Оператор вывода	Общее выражение
26.	Оператор разметки	Арифметическое выражение
27.	Совокупность спецификаций	-
28.	Ограничитель параметра	-
29.	Стандартная функция	Арифметическое выражение

В таблице № 1 для каждого из вызываемых распознавателей приведен список соответствующих вызываемых распознавателей, а также для некоторых из них указаны контролеры ошибок. При обнаружении ошибки распознаватель выдает соответствующую информацию (см. таблицу № 2) и в определенных случаях подключает контролеры ошибок. Последние учитывают возможные ошибочные ситуации, обеспечивают нейтрализацию опознанных ошибок распознавателями, а также определяют возможность продолжения анализа. Таким образом, процесс передачи управления распознавателями осуществляется рекурсивно.

Таблица № 2 иллюстрирует функционирование распознавателя арифметического выражения.

Таблица № 2

Текущее состояние	Символ	Действие	Следующее состояние	
			нет ошибок	ошибочное состояние
1	2	3	4	5
АРИФ1	- +	Выбор следующего символа	АРИФ2	
	<i>if</i>	Выбор следующего символа, распознаватель булевского выражения	АРИФ5	АРИФ9
АРИФ2	Число	Выбор следующего символа	АРИФ4	
	Идентификатор	Получаем тип идентификатора	АРИФ8	
	(Выбор следующего символа, распознаватель арифметического выражения	АРИФ3	АРИФ10
	другое	Сообщение об ошибке, контролер арифметического выражения		Ошибочный возврат
АРИФ3)	Выбор следующего символа	АРИФ4	
	другое	Сообщение об ошибке.	АРИФ10	
АРИФ4	арифмет. операция	Выбор следующего символа	АРИФ2	
	другое		Возврат	
АРИФ5	<i>then</i>	Выбор следующего символа	АРИФ6	
	другое	Сообщение об ошибке	АРИФ9	
АРИФ6	<i>if</i>	Сообщение об ошибке, распознаватель арифметического выражения	АРИФ7	АРИФ9
	другое	Распознаватель арифметического выражения	АРИФ7	АРИФ9
АРИФ7	<i>else</i>	Выбор следующего символа, распознаватель арифметического выражения	Возврат	Ошибочный возврат
	другое	Сообщение об ошибке	АРИФ9	
АРИФ8	Переменная типа <i>real</i> или <i>integer</i>	Выбор следующего символа	АРИФ4	

1	2	3	4	5
	<i>real procedure</i> / <i>integer procedure</i>	Выбор следующего символа, распознаватель функции	АРИФ4	Ошибочный возврат
	<i>array</i> или <i>real array</i> / <i>integer array</i>	Выбор следующего символа, распознаватель индексного выражения	АРИФ4	Ошибочный возврат
	переменная типа <i>Boolean</i> / <i>Boolean array</i>	Сообщения об ошибке, выбор следующего символа	АРИФ4	
	<i>Boolean procedure</i>	Сообщение об ошибке, выбор следующего символа, распознаватель индексного выражения	АРИФ4	Ошибочный возврат
	Стандартная функция	Сообщение об ошибке, выбор следующего символа, распознаватель функции	АРИФ4	Ошибочный возврат
	другое	Расознаватель стандартной функции	АРИФ4	Ошибочный возврат
АРИФ9	<i>then else</i>	Сообщение об ошибке, контролер арифметического выражения	АРИФ5 АРИФ7	Ошибочный возврат
АРИФ10)	Контролер арифметического выражения	АРИФ3	Ошибочный возврат
	другое	Контролер арифметического выражения		Ошибочный возврат

Контроль программы осуществляется по схеме, показанной на рис. I.

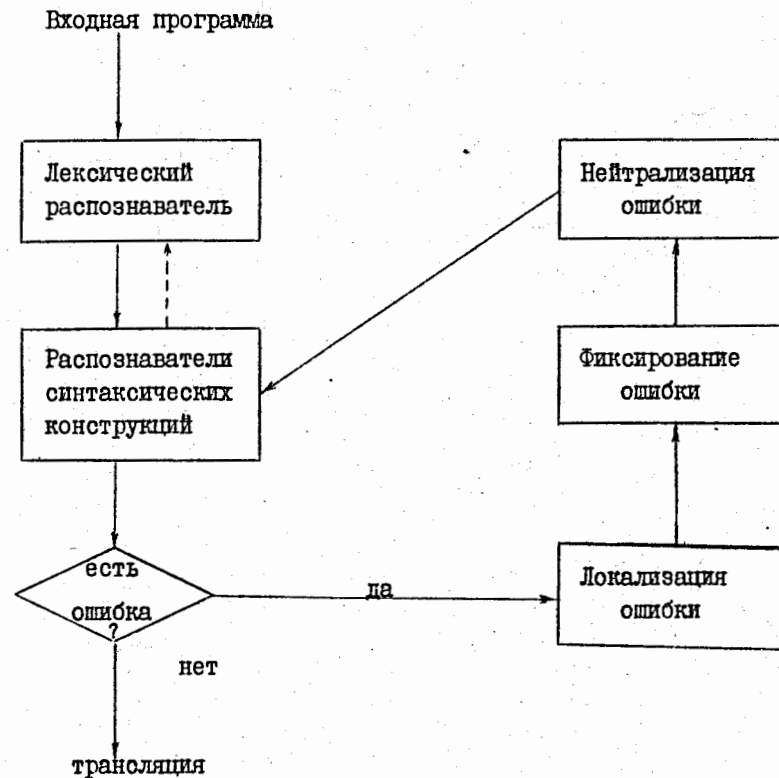


Рис. I. Функциональная блок-схема контроля

В рассматриваемой системе отладки обеспечивается пропуск минимальной цепочки символов. Сообщения об ошибках классифицируются на предупреждающие - о возможных ошибках, корректирующие - об устранимых системой отладки ошибках, сообщения о грубых ошибках, об ошибках с прекращением контроля и выдаются в удобном для понимания виде.

Предложенный в работе метод распознавателей осуществляет глубокий контроль исходных программ. Так в тестовом примере, рекомендованном Кнутом и содержащем 25 ошибок, транслятор АЛГАМС ДОС ЕС ЭВМ за один просмотр обнаружил 20 ошибок (тогда как другие известные трансляторы в том же тестовом примере обнаруживают за один просмотр не более 12 ошибок (см. приложение № 1 к диссертации)).

В приложении № 2 к диссертации приводятся примеры программ, содержащих типичные ошибки с информацией, выдаваемой о них транслятором ТАМ-32.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Непосредственно автором получены следующие результаты, отраженные в диссертации:

- предложен и разработан метод полного синтаксического и частичного семантического контроля программ - метод распознавателей, ориентированный на класс языков типа АЛГОЛ и обеспечивающий нейтрализацию и глубокую локализацию обнаруживаемых ошибок;
- предложенный метод распознавателей реализован и внедрен в ряде производственных систем отладки программ в трансляторах ТАМ-2(22), ТАМ-22Т, АЛГАМС ДОС ЕС ЭВМ непосредственно автором и в трансляторе ТАМ-32.

Основное содержание диссертации нашло отражение в следующих работах:

1. Л.И.Бородич, Л.Б.Куракина, О контроле АЛГАМС-программ, сб. "Математическое обеспечение ЭВМ "Минск-32", вып. 14, Минск, 1975, 128-141.

2. Л.И.Бородич, Л.Б.Куракина, Структура анализатора АЛГАМС-программ, сб. "Математическое обеспечение ЭВМ "Минск-32", вып. 15, Минск, 255-262.

3. Л.И.Бородич, Л.Б.Куракина, Структура АЛГАМС-программы и ее связь с таблицами синтаксического разбора, сб. "Математическое обеспечение ЭВМ "Минск-32", вып. 15, Минск, 1975, 263-271.

4. Л.И.Бородич, О синтаксическом контроле и способе локализации ошибок в АЛГАМС-программах, сб. "Проблемы развития прикладных математических исследований", Тезисы докладов IV Республиканской конференции математиков Белоруссии, Минск, 1975, 84-85.

5. Л.И.Бородич, Г.П.Свирид, Защита памяти при адресации компонент массивов, сб. "Математическое обеспечение ЭВМ "Минск-2(22)" в режиме Т, вып. 10, Минск, 1972, 88-89.

6. Л.И.Бородич, Р.Ф.Дегтярева, Н.В.Шкут, К вопросу автоматизации программирования задач численного анализа для машин типа "Минск", сб. Материалы республиканской научно-технической конференции "Применение математических методов и вычислительной техники в народном хозяйстве", Минск, 1968, 32-33.

7. Л.И.Бородич, О.Н.Жевняк, Н.К.Касько, Л.И.Кузьмина, В.К.Сидорик, Н.Н.Черп, Н.В.Шкут, АЛГАМС в системе ДОС ЕС ЭВМ, "Статистика", М., 1977, 3-III.

8. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, Н.К.Касько, Л.И.Кузьмина, О.В.Подольская, В.К.Сидорик, Н.Н.Черп, Н.В.Шкут, Реализация языка АЛГАМС для ЕС ЭВМ (операционная система ДОС), сб. "Проблемы развития прикладных математических исследований", Тезисы докладов IV Республиканской конференции математиков Белоруссии, Минск, 1975, 85-86.

9. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, С.В.Попова, Т.С.Шинкевич, Н.В.Шкут, Транслятор ТАМ-2, Республиканский фонд алгоритмов и программ, Минск, 1968.

10. Л.И.Бородич, Р.Ф.Дегтярева, О.Н.Жевняк, Транслятор ТАМ-22, Республиканский фонд алгоритмов и программ, Минск, 1968.

11. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, С.В.Попова, Г.П.Свирид, И.Г.Трухан, Г.С.Шинкевич, Н.В.Шкут, О трансляторе ТАМ-22Т, сб. "Всесоюзная межвузовская конференция по математическому обеспечению автоматизированных систем управления", Тезисы докладов, Москва, 1970, 26-28.

12. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, С.В.Попова, Г.С.Шинкевич, Н.В.Шкут, Транслятор ТАМ-22Т, Республиканский фонд алгоритмов и программ, Минск, 1971.

13. Л.И.Бородич, Г.П.Свирид, О трансляции АЛГАМС-программ, сб. "Математическое обеспечение ЭВМ "Минск-2(22)" в режиме Т", вып. 8, Минск, 1971, 16-30.

14. Л.И.Бородич, О.Н.Жевняк, Н.К.Касько, Л.И.Кузьмина, В.К.Сидорик, Н.Н.Черп, Н.В.Шкут, Руководство по использованию АЛГАМСа в ДОС ЕС ЭВМ, сб. "Математическое обеспечение ЕС ЭВМ", вып. 8, Минск, 3-203.

15. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, С.В.Попова, Г.С.Шинкевич, Н.В.Шкут, О некоторых особенностях транслятора ТАМ-22Т, сб. "Третья республиканская конференция математиков Белоруссии", Тезисы докладов, часть II, Минск, 1971, 171.

16. Л.И.Бородич, Р.Ф.Дегтярева, З.А.Единович, О.Н.Жевняк, С.В.Попова, Г.С.Шинкевич, К вопросу эксплуатации ТАМ-22Т, сб. "Математическое обеспечение ЭВМ "Минск-2(22)" в режиме Т", вып. 10, Минск, 1972, 78-92.

17. Л.И.Бородич, З.А.Единович, С.В.Попова, Трансляторы ТАМ-2 и ТАМ-22 (руководство по эксплуатации), "Наука и техника", Минск, 1967, 1-13.

АТ 13524. Подписано в печать 10.03.1978г.
Формат 60x84 1/16. 0,93 усл. печ. л., 0,68 уч. изд. л.,

Тираж 100 экз. Заказ 47.

Отпечатано на ротационной машине Института математики АН БССР,
г. Минск, ул. Типографская, 11.