

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА



Ц 8405

К-672

27/x-

10 - 9126

А.А. Корнейчук

У150/2-75

СИСТЕМА АСТРА.

ЛЕКСИЧЕСКИЙ АНАЛИЗАТОР
АВТОКОДНОГО ТЕКСТА

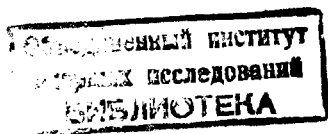
1975

10 - 9126

А.А.Корнейчук

СИСТЕМА АСТРА.

ЛЕКСИЧЕСКИЙ АНАЛИЗАТОР
АВТОКОДНОГО ТЕКСТА



АСТРА, система автоматизации программирования для ЭВМ БЭСМ-4, разработана в Лаборатории вычислительной техники и автоматизации Объединенного института ядерных исследований (Дубна).

Версия 74/1 системы АСТРА (автокод, комплектование библиотеки стандартных подпрограмм, редактирование) описана в работах /1-4/. Со времени опубликования этих работ система АСТРА пополнилась рядом новых программ. В работе /5/ описан транслятор с макрокода. Настоящее сообщение посвящено лексическому анализатору автокодного текста. Анализатор написан на автокоде и снабжен подробными пояснениями, что позволяет, при необходимости, вносить изменения и использовать отдельные блоки в других программах.

Назначение лексического анализатора

Назначение лексического анализатора автокодного текста - распознавание лексических единиц (меток, кодов операций, идентификаторов и др.) и приведение текста к формату (будем называть его нормальным, или N-форматом), пригодному для работы транслятора с автокода АСТРА (см. /1-2/). В версии 74/1 автокода лексический анализатор отсутствовал и автокодный текст записывался в N-формате. Известным неудобством N-формата является нерациональное использование поля перфокарты (мало места для пояснений). В связи с этим введен новый формат автокодного текста (будем называть его свободным, или F-форматом). Поскольку N-формат обладает, тем не менее, определенными достоинствами (например, большей наглядностью), то сохранена возможность произвольного смешивания перфокарт, пробитых в N- и F-форматах.

Стандартная подпрограмма лексического анализа

Лексический анализатор, или нормализатор автокодного текста (NAT) - стандартная подпрограмма в системе ИС-2. Обращение к СП NAT:

016 +001 7501 7610

052 ВТХТ NAT ЕТХТ

ВТХТ и ЕТХТ - начало и конец нормализуемого текста. Нормализованный текст помещается на место исходного.

Лексический анализатор может быть использован для предварительной нормализации текста, хранящегося на магнитной ленте, а также непосредственно в составе транслятора с автокода.

Распознавание карт в N- и F-форматах

Карта в F-формате (F-карта) начинается с одного пробела. Среди следующих 5 символов F-карты должен быть хотя бы один, отличный от пробела.

Карта в N-формате (N-карта) обязана начинаться либо с 6 пробелов, либо с символа, отличного от пробела.

Особенности F-карт

1. В F-картах, не являющихся комментариями, пробелы могут быть использованы произвольно — при нормализации они игнорируются.

2. Метка отделяется от кода операции, кода псевдокоманды, текстовой константы либо десятичной константы запятой. Например:

L, 501, A, A+1, B, сложение A и A+1

, 112, N1, -1, 1, конец цикла

3. Код операции отделяется от первого адреса запятой. Стоящие в начале кода операции нули могут быть опущены.

4. Код псевдокоманды отделяется от адреса запятой.

5. Первый адрес (в команде) отделяется от второго адреса запятой.

6. Второй адрес (в команде) отделяется от третьего адреса запятой.

7. Третий адрес (в команде) отделяется от текста (содержащего пояснения) запятой.

8. Адрес (в псевдокоманде) отделяется от текста (содержащего пояснения) запятой.

9. Запятые, перечисленные в пунктах 2-8, могут быть опущены, если следующие за ними элементы — коды операций, коды псевдокоманд, адреса, тексты, содержащие пояснения, — состоят из пробелов.

10. В управляющих комментариях перед начальной звездочкой может быть 1-5 пробелов.

11. В текстовой константе длина текста отделяется от следующего за ней текста запятой. Стоящие в начале длины текста нули опускать нельзя.

12. В десятичной константе порядок отделяется от мантиссы запятой. Знак числа и знак порядка опускать нельзя. Стоящие в начале порядка нули можно опустить.

13. Мантисса десятичной константы отделяется от текста, содержащего пояснения, запятой. В мантиссе допустимы десятичные цифры (не более 9) и пробелы. При нормализации пробелы игнорируются. Недостающие до 9 последние цифры полагаются равными нулю.

14. Нули, стоящие в начале метки, могут быть опущены.

15. Нули, стоящие в начале истинного адреса, могут быть опущены.

16. Нули, стоящие между знаком "+" или "-" и первой отличной от нуля восьмеричной цифрой относительного адреса, могут быть опущены.

17. В символических адресах, содержащих более 6 символов, остаются только первые 6 символов.

Выдача сообщений об ошибках

Если лексический анализатор обнаруживает ошибку, на АЦПУ печатается сообщение, в котором указан условный номер ошибки, слово в карте и номер карты в массиве текста (в восьмеричной системе, начиная с нуля), в которой была обнаружена ошибка. Значение условных номеров ошибок:

1 — истинный адрес не состоит из 1, 2, 3 либо 4 восьмеричных цифр;

2 — ошибка в порядке десятичной константы;

3 — в мантиссе десятичной константы встретился символ, не являющийся пробелом либо десятичной цифрой.

О тексте лексического анализатора

В тексте СП, приведенном в настоящем сообщении, вместо звездочки использован апостроф, Автокодные предложения, помещенные для экономии места в одну строку, разделяются двоеточием.

Заключение

Автор благодарит Л. П. Корнейчук за помощь при подготовке рукописи к печати.

Д1,2-7411	Глубоконеупругие множественные процессы. Дубна, 1973.	507 стр.	5 р. 66 к.
Д13-7616	Труды VII Международного симпозиума по ядерной электронике. Будапешт, 1973.	372 стр.	3 р. 65 к.
Р1,2-7642	Труды Международной школы молодых ученых по физике высоких энергий. Гомель, 1973.	623 стр.	7 р. 15 к.
Д10-7707	Совещание по программированию в математических методах решения физических задач. Дубна, 1973.	564 стр.	5 р. 57 к.
Д1,2-7781	Труды III Международного симпозиума по физике высоких энергий и элементарных частиц. Сивая, 1973.	478 стр.	4 р. 78 к.
Д3-7991	Труды II Международной школы по нейтронной физике. Алушта, 1974.	552 стр.	2 р. 50 к.
Д1,2-8405	Труды IV Международного симпозиума по физике высоких энергий и элементарных частиц. Варна, 1974.	376 стр.	2 р. 05 к.
Д10,11-8450	Труды Международной школы по вопросам использования ЭВМ в ядерных исследованиях. Ташкент, 1974.	465 стр.	2 р. 46 к.
Р1,2-8529	Труды Международной школы-семинара молодых ученых. Актуальные проблемы физики элементарных частиц. Сочи, 1974.	582 стр.	2 р. 60 к.

Заказы на упомянутые книги могут быть направлены по адресу:

101000 Москва, Главпочтамт, п/я 79,

издательский отдел Объединенного института ядерных исследований.