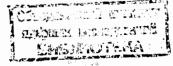


А.И. Волков

ABTOКОД MADLEN (ОПИСАНИЕ ТРАНСЛЯТОРА)



Введение

В системе программирования ДУБНА /4/ автокод марсем применяется в случаях, когда при программировании необходимо максимально учитывать особенности машины БЭСМ-6, а также в качестве промежуточного языка при трансляции с ФОРТРАНа. В будущем, возможно, автокод марсем используют трансляторы с каких-либо других алгоритмических языков. Таким образом, эффективность трансляции в системе существенно зависит от скорости трансляции с автокода. Это обстоятальство повлияло на структуру транслятора с автокода, предусматривающего различные варианты работы в зависимости от "ти-па" пользователя.

Так как транслятор с автокода написан на автокоде МАДLEN и текст транслятора снабжен подробными комментариями, данное описание не ориентировано на подробное изложение всех деталей. В сообщении описаны лишь принципиальные вопросы построения транслятора и приводятся некоторые сведения, упрощающие непосредственное ознакомление с текстом транслятора. В приложениях, кроме того, приведена инструкция к подпрограмме транслятора SIMFAIL*, предназначенной для работы с последовательностями (лентами) текстовых восымиразрядных символов. Эта подпрограмма может оказаться полезной для применений, не связанных с рассматриваемым транслятором.

Предполагается, что читатель имеет общее представление о структуре мониторной системы ДУБНА, знаком с описанием автокода MADLEN.

І. Обращение к транслятору

Обращение к транслятору имеет вид:

- I, VTM , READ/
- 2, VTM , SAVE/
- 3, VTM , MADLEN/
- 4, VTM , 4096
- 5, VTM , TØ
 - , CALL , MADLEN

- $T\emptyset$ максимальная длина таблицы описаний (и поля, отведенного для ее формирования). В настоящем варианте $T\emptyset$ = IO24.
- READ/ подпрограмма выдачи очередной инструкции транслируемой подпрограммы. Инструкция всегда содержит I4 слов.
 READ/ удовлетворяет библиотечным правилам и имеет один формальный параметр адрес размещения текста инструкции.
- SAVE / подпрограмма запоминания и восстановления состояния READ / . Используется для организации двух просмотров текста исходной подпрограммы. Единственный параметр задается значением сумматора перед обращением к SAVE/:

Σ = 0 - запоминание $Σ \neq 0 -$ восстановление.

Возможны две формы задания инструкции (см. 2):

- а) текстовая карта (84 символа ISØ),
- б) набор понятий (слов)

 $MADLEN/ = \begin{cases} MADLEN* - работа с текстовыми картами \\ MADLEN** - работа с наборами понятий \\ MADLEN** - подпрограммы автокода$

- 2. Две формы инструкций
- I. <u>Текстовые карты</u>. Текстовой картой нашывается последовательность символов

символы 150/3/, а $S_{84}=0/2_8$ — признак конца последовательности.

Текстовые карты являются общеупотребительным способом обмена текстами между подпрограммами монитора, к которым следует отнести и трансляторы.

Текстовые карты используются при трансляции подпрограмы непосредственно с автокода (активизируемого в этом случае монитором).
В качестве фактических параметров, соответствующих формальным параметрам READ/ и SAVE/, используются наименования подпрограмы
постоянной библиотеки.

2. <u>Набор понятий</u>. В случае, когда программа на автокоде является результатом работы транслятора с ФОРТРАНа, нет необходимости формировать обычный текст подпрограммы, требующий трудоемкой работы по выделению основных синтаксических понятий, так как таковые уже были определены на уровне трансляции с ФОРТРАНа.

Набор понятий, характеризующий одну инструкцию автокода, определяется группой из семи слов:

- $\mathcal{R}_{c} = I$ Содержимое набора игнорируется. На печать (в режиме FULL LIST) выдается строка С ****** НАЧАЛО ОПЕРАТОРА $_{L}$ $_{C}$ $_{C}$
- $\pi_{I} = I$ Первая составляющая адреса (предполагаемая всегда положительной) есть идентификатор.
- $\pi_L = I$ Первая составляющая адреса адрес типа "ЛИТЕРАЛ", причем соответствующая константа состоит из одново слова, если $\pi_D = Q$.
- $(\pi_{L^{\pm}1})$ π_{D} = I Конатанта, соответствующая конструкции "адрес типа литерал", состоит из двух слов.
- 2. метка в коде Т Е X Т

 48 4 I

 л_м мер

ж_м= I - Инструкция помечена.

мор - определяет значение указателя регистра в инструкции.

5.

Последние два слова используются для представления первой составляющей адреса:

- а) Если \mathcal{R}_{I} = I или мнемокод есть вьюск (или сомт) в A_{I} помещается идентификатор в коде T Е X T. AI игнорируется.
- б) Если $\mathcal{I}_L = I$ ($\mathcal{I}_T = 0$), то соответствующая константа помещается в AI (при $\mathcal{I}_D = 0$) или же в AI и AI' ($\mathcal{I}_D = I$)
- в) Если мнемокод есть Ø C T (другие константы недопустимы), то константа помещается в AI.

A2 - вторая составляющая (00000+77777₈)

 $\pi_{\pm} = 0 +$ $\pi_{\pm} = I -$ знак перед второй составляющей

Для \mathcal{BLQCK} - вторая составляющая есть длина фрагмента (в адрес представим только один фрагмент). Состояние $\mathcal{H}_{\mathbf{Z}}$ безразлично.

Нетрудно заметить, что в данном случае на синтаксис автокода накладывается ряд ограничений: инструкция выбск имеет только один фрагмент в адресе, значение указателя регистра не может быть идентификатором и т.п. При необходимости эти ограничения могут быть сняты посредством увеличения понятий в наборе. Данный вариант набора определен спецификой трансляцыи с языка ФОРТРАН в системе ДУБНА.

Обращение к транслятору в этом случае осуществляется заключительным блоком транслятора с ФОРТРАНа (GLINK) /I/ с параметрами READF, SAVEF и MADLEN**, соответствующими формальным параметрам READ/, SAVE/ и MADLEN/

3. Состав транслятора

Приведенная ниже **схема иллострирует** взаимодействие подпрограмм автокода и некоторых других подпрограмм (из постоянной библиотеки), используемых при трансляции с автокода:

ASEMBLER - заполняет признак "трансляция непосредственно с ав-

токода" (MADLEN*). Выполняет обращение к транслятору с параметрами

READ*, SAVE*, MADLEN*.
Обращение к этой подпрограмме осуществияется из монитора посредством

, CALL , ASEMBLER

- марыем начальная подпрограмма транслятора. Выполниет подготовку к трансляции.
- .марьем блок трансляции непосредственно с автокода.

 Транслируемая подпрограмма последовательность текстовых карт.
- мариєм** блок трансляции с автокода. Транслируемая подпрограмма - последовательность наобров понятий

READ *
SAVE *
READF
SAVE F

описаны выше.

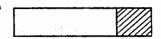
- корнаме таблица допустимых в марьем мнемонических кодов.
- PRIMADL* выполняет сервисные работы при трансляции (печать, редактирование и т.п.).

nep6881C BCHOMOTATE

вспомогательные подпрограммы для PRIMADL

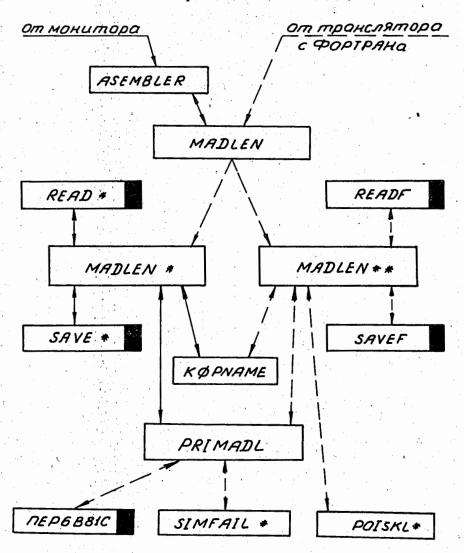
POISKL* — часть MADLEN**. Выделена в отдельную подпрограмму для использования ее и транслятором с $\Phi OPTPAHa$ (блок GLINK).

Пунктирными линиями на схеме связаны подпрограммы, работающие только при трансляции с ФОРТРАНа. Таким образом, различным вариантам использования транслятора с автокода соответствуют различные варианты загрузки его подпрограмм. Обозначение блока



использовано для подпрограми, не относящихся непосредственно к комплексу транслятора с автокода.

<u>Схема I.</u> Взаимодействие подпрограми при трансляции с автокода



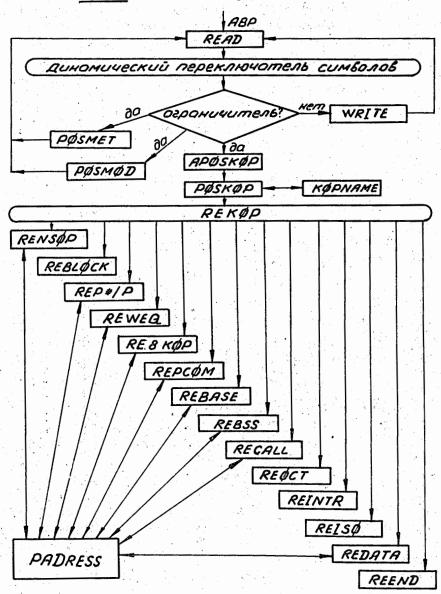
4. Схема трансляции

Трансляция осуществляется прямым (или несинтаксычески, спосопроводится два просмотра текста транслирусмой oom). Hou stom подпрограммы. На первом просмотре определяются некоторые Количественные характеристики (структура) стандартного массива, выполняется полная обработка меток, некоторых эквивалентностей, а также ряд операций вспомогательного характера. В результате работы транслятора на первом просмотре формируется справочная таблица, содержацая сведения о характере всех объектов, использованных в транслируемой подпрограмме (массивы, константы типа "литеран", подпрогрешмы и т.п.). Эта таблица размещена в массиве Sekt@R . Будсы ее далее называть условно таблицей идентирикаторов. На втором просмотре окончательно формируются все элементы стандартного массива и выполняется, если необходимо, нечать и перфорация результатов трансляции.

Схема 2 иллюстрирует разоту транслятора при выполнении одного просмотра блоком трансляции марсем . Одник из основных рунклиональных элементов блока является динамический переключатель символов, параметр которого есть код очередного символа транслируемой подпрограммы, поступающего от процедуры кеар . Совокупность символов, расположенных в исходной подпрограмме между двумя последовательными ограничителями, запоминается процедурой write и в зависимости от типов окаймляющих ограничителей и состояния переключателей передается одной из подпрограмм РØS мет , РØS мФD или
АРØS кФР , завершающих, соответственно, обработку указателя метки, указателя регистра или мнемонического кода операции.

Если мнемокод не является заголовком подпрограммы (этот случай обрабатывается самой подпрограммой *АРОККОР*), то он передается подпрограмме *РОККОР* для окончательной обработки. С помощью таблицы мнемокодов здесь определяется параметр переключателя *RЕКОР*, выполняющего передачи управления на процедуры обработки конкретных типов инструкций. Эти процедуры, в свою очередь, могут потребовать выделения понятил из адресной части инструкции. Для большинства типов инструкций адрес имеет форму полного адреса, обработкой которого знимается *PADRESS*. Для других инструкций рассмотрение адреса выполняется непосредственно процедурами обработки конкрет-

<u>Схема 2</u>. Работа блока трансляции MADLEN ★



ной инструкции. Однако почти во всех случаях выделение понятий из адресной печати выполняется с помощью уже описанного аппарата (динамического переключателя символов и процедур READ и WRITE).

Схема работы блока трансляции марсем** аналогична описанной выше. Однако эта подпрограмма не содержит элементов, связанных с выделением понятий из исходного текста. марсем** не имеет динами-ческого переключателя, процедур READ и write. В серию других процедур также внесены соответствующие исправления, обусловленные отсутствием работы по выделению понятий, а также упрощениями в синтаксисе входного языка.

5. Используемые массивы

Транслятор использует следующие массивы:

а) для связи с монитором:

FULIS FL # - признак режима FULL LIST

NOLISFL * - признак режима NO LIST

маDLE № * - признак трансляции непосредственно с автонода;

ENTRY* - таблица входов в подпрограмму;

FL6EX * - характеристика ошибок трансляции.

б) собственные массивы транслятора:

RABINF (200) - константы и некоторые рабочие массивы;

SECTOR (4096) - таблица идентификаторов;

TABDEC (1024) - таблица описаний;

ERRORBUF(13) - рабочий массив процедур;

SPASAN* (210) - рабочий массив, используемый процедурами

REISØ (MADLEN*), EDITØR и EDITØR F (PRIMADL*);

ВUFPER (256) - буфер для записи элементов стандартного массива во временную библиотеку.

RABINF

Список всех констант, помещенных в массив *PABINF*, приведен в подпрограмме *MADLEN*. В приложении I дан перечень наиболее часто употребляемых элементов этого массива (рабочих ячеек) с указанием специфики их использования. Поэтому ниже приводится лишь описание двух элементов этого массива (*MISTABLE* и *FLAGTABL*), используемых для управления трансляцией.

мізтавь — таблица локальных признаков трансляции. Такие признаки сохраняются, как правило, лишь на период трансляции одной инструкции.

48 47 46 45 44 43 24 23 22 16		T					 _		_			
	48	47	46	45	44	43	 24	23	22			
		1 '''	, , ,		'''					6	+ -	

Здесь

- 48 р. признак ошибки в инструкции;
- 47 р. признак одибки в метке;
- 46 р. признак ошибки в указателе регистра;
- 45 р. признак ошибки в мнемокоде;
- 44 р. признак ошибки в адресе;
- 24 р. признак помеченности инструкции:
- 23 р. признак команды с длинным адресом;
- 22 р. признак "модификатор-идентификатор";
- 16 р. признак "недопустим нулевой модификатор".

FLAGTABL - таблица глобальных признаков трансляции.

46	1 1	43	24 2	23 22	5	I
<u> </u>						لــــــــــــــــــــــــــــــــــــــ

- 46 р. признак групп DATA и set
- 43 р. признак группы ДАТА
- 24 р. признак помеченности следующей команды;
- 23 р. признак переноса помеченности по требованию предыдущей команды;
 - 5 р. признак трансля им с автокода *марьем* (если 0, то трансляция с sibes м -6);
 - I р. признак первого просмотра текста транслируемой подпрограммы.

SECTOR

Таолица идентификаторов содержит сведения обо всех идентификаторах (за исключением дополнительных входов) и константах, определяемых конструкциями "адрес типа литерал". Каждому такому объекту соответствуют два слова:

- I) идентификатор в коде TEXT или константа,
- 2) характеристика.

Характеристика для идентификаторов имеет следующее строение:

29...19 18.17 16 A 45 р. - признак неопределенности описания; 44 р. - признак двойственного описания; \forall = I - werka: $\nu = 3$ - подпрограмма; v=3.46 р.=I. — подпрограмма, имеющая описатель У = 4.7 и У определяют массивы так же, как в таблице описаний: А - значение идентификатора: I6 р. - признак условного A: V = 0 - эквивалентность: 43 р. = 0 - истинному адресу или метке в зависимости от значения 16 р. 43 р. = I - относительному адресу А для констант "литерал" блока MADLEN* : 48.... 46 A V = I -константа типа LOG ; V = 2 - тип INT (указатель I); (указатель R): $V = 3 - TM\Pi REAL$ V = 4 - тип ISØ (указатель Н);А - ссответствурний относительный адрес. для констант "литерал" блока MADLEN++: A' V = I -константа из одного слова ($A^I = 0$): $\gamma = 2$ - первое слово двойной константы: A^{\perp} - адрес в таблице второго слова; $\sqrt{3}$ = 3 - второе слово двойной константы (A=A¹=0);

А - относительный адрес, соответствующий константе.

6. Некоторые детали

- а) <u>лепользование регистров</u>. К числу глобальных индексных регистров относены регистры I,3 7. Основное их назначение определяется следующим списком:
 - I начальным адрес динамического переключателя:
 - 3 м 4 для работы процедуры READ (MADLEN *);
 - 5 для работы процедуры WRITE (MADLEN);
 - 6 STRØKA :
 - 7 база констант процедуры PADRESS (MADLEN *).
- б) $\frac{5}{2}$ нкция расстановки. Для работы с таблицей идентификаторов применяется аппарат бункций расстановки, описанный в $\frac{2}{2}$. Применяется следующий вариант функции расстановки: идентификатор (или константа) режется на куски длиной по II разрядов, которые складываются по мод 2^{II} . Обозначим результат через s. Тогда функция расстановки определяется формулой
- f (<идентификатор>) = A + S * 2, где A начальный адрес функции расстановки.
- в) Подпрограмма SIMFAIL*. Эта подпрограмма приводится в приложении 2 и может быть использована для различных текстовых преобравований. Так как при создании подпрограммы основное внимание
 уделялось ее универсальности, в ряде случаев ее применение может
 привести к написанию недостаточно быстродействующих программ. Так,
 в подпрограмме MADLEN* процедуры для выделения понятий READ и
 WRITE используют более эффективный частный алгоритм.
- г) Транслиция с автокода SIBESM-6 /5/. Совместимость с автокодом SIBESM-6 ставилась лишь как частная задача по сохранению без значительных изменений программ, написанных на этом языке для версии транслитора, используемого в течение 1969 года. Для возможных последующих версий такой совместимости может и не быть. Что же касается "совместимого" варианта SIBESM-6, то для его трансляции требуются следующие изменения: инструкции DATA, SET, END должны быть написаны в соответствии с синтаксисом автокода марьем.

д) Характеристики транслятора:

Объем транслятора при трансляции непосредственно с автокода — ≈ 4000 слов;

Объем транслятора при использовании

набора понятий — ≈ 2500 слов;

Объем рабочих полей — ≈ 3000 слов;

Максимальное количество объектов

в подпрограмме (идентификаторов

и констант "литерал") - 2048;

Скорость трансляции непосредственно

с автокода - около 200 инструкций в секунду;

Скорость трансляции при использовании наборов понятий

около 400 инструкций в секунду.

ПРИЛОЖЕНИЕ I Maccub RABINF

RABINF : LC , BLØCK , (10) , STRØKA (190)

(IO) - вспомогательное поле для организации печати очередной инструкции (для хранения инструкции отведено I4 первых слов массива STRØKA);

АА - начало таблицы идентификаторов;

ААА - начало таблицы описаний;

ABSCNT - КОЛИЧЕСТВО АБСОЛЮТНЫХ КОНСТАНТ;

AENTR - начало таблицы входов;

ВВ - конец таблицы идентификаторов;

ВВВ - конец таблицы описаний;

ВАЅЕМОД - значение указателя регистра в инструкции;

BASE* - значение (- БАЗА) при заданном автоматическом базировании;

BASIR - номер регистра, по которому должно производиться автоматическое базирование;

BASPR - признак "инструкция-команда";

ВЕ NTR - конец таблицы входов;

BSSCNT - количество последних BSS;

DATENT - КОЛИЧЕСТВО КОНСТАНТ **DATA**

FLAGTABL - таблица общих признаков трансляции (см. 5);

INDTABL - padovee none процедуры WRITE (MADLEN*)

IMIST - признак ошибки при выделении понятия (или его синтаксическом анализе);

козсит - количество констант "литерал";

LEFINST - ячейка для формирования очередного слова из двух команд;

мЕSADR1 - первая составляющая полного адреса;

MESADR 2 - вторая составляющая полного адреса;

MESAD R3 - значение полного адреса в случае истинности;

MESKOP - MHEMOKOZ;

MESMET - METHA;

мезмøр - указатель регистра; /

MESREAD * - фактический параметр для READ/

MESSAVE * - фактический параметр для SAVE/

MISTABL - таблица локальных признаков трансляции (см. 5);

NKABS - начальный адрес группы абсолютных констант;

NKCARD - порядковый номер карты подпрограммы;

NKKØS - начальный адрес группы констант "литерал";

ØUTMADLE - возврат из автокода;

POSIND - понятие, выделенное процедурой write;

RMIST - признак ошибки при переводе понятий из одной системы представления в другую.

<u>ПРИЛОЖЕНИЕ 2.</u> Подпрограмма для работы с лентами символов З*IMFAIL* ► Назовем лентой символов последовательность

S₁, S₂, ..., S_N,

где S_t - восьмиравные символы (000+3778), а N- кратно 6-ти. Предполагается, что такая последовательность плотно закодирована (в порядке следования символов) в массив ячеек оперативной памятидлиной $\frac{N}{6}$. Поставим в соответствие такой ленте указатель j значение которого является индексом интересующего нас символа S_j

 $1 \le i \le N$. Для такой ленты определим следующие операции:

I. Запись символа в ленту (для установленного к этому моменту значения указателя).

- 2. Чтение символа с ленты.
- 3. Увеличение указателя на единицу (операция определена для $j \leq N-1$).
- 4. Уменьшение указателя на единицу (операция определена для $j \gg 2$).

Так как программа предназначена для работы с несколькими лентами, то, в общем случае, ленту определяют следующие параметры:

А - начальный адрес ленты,

і - указатель ленты,

 $n = \frac{N}{6}$ - количество слов, отведенных в памяти для ленты.

Для хранения параметров каждой ленты отводится характеристи— ка X — массив из четырех слов.

Подпрограмма SIMFAIL* является набором процедур, реализующих различные операции с лентой. Эти процедуры могут быть описаны следующим образом:

SIMFAIL *: , SUBP,
SIMFAIL *: I, BLØCK, DØ * ЛЕНТ, ПЛЮСІ
, CØNT , MUHYCI, ЗАП * ПІ
, CØNT , 4TE * ПІ, ПІ * ЗАП
, CØNT, ПІ * 4TE, МІ * ЗАП
, CØNT, МІ * 4TE, ЛІ * ТØ * Л2
, CØNT, 4TEHUE, ЗАПИСЬ
, CØNT, Л2 * ТØ * Л1

Тогда, например, обращение к процедуре ЧТЕНИЕ осуществляется следующим образом:

13, VOM, TEHNE

Входные параметры

18 - X

19 - A

I10 - n

Здесь и ниже обозначение $\mathbf{I}_{\mathbf{K}}$ применяется для указания содержимого индексного регистра с номером К.

11 * 10 * 12 перепись характеристики X1 в массив X2

18 - X1,

19 - X2

12 * 70 * 11 перепись характеристики X2 в массив X1:

18 - X1

I9 - X2

Все другие процедуры реализуют операции для ленты, <u>жарактари-</u> стика которой задана в 18.

Элементарные операции:

ЧТЕНИЕ - чтение символа. Символ помещается в 114.

ЗАПИСЬ - запись символа. Исходный символ должен быть помещен в 114.

ПЛЮС 1 — увеличение указателя на единицу. Если в результате i = N+1 , то I11 $\neq 0$.

МИНУС1 — уменьшение указателя на единицу. Если в результате i=0, то I11 $\neq 0$.

Все остальные процедуры реализуют композиции тех или иных элементарных операций. Так, например, инструкция

13 , VJM , ЗАП * П1

эквивалентна следующей паре инструкций

13 , VJM , BATHCL

13 , VJM, IIJDC1

Все подпрограмми сохраняют значения I1 +I8. Из остальных регистров каждая процедура гарантирует сохранение лишь тех, которые используются для задания информации к ней.

литература

- Бродцински З., Веретенов В.Ю. и др. Транслитор с языка ФОРТРАН для системы математического обеспечения БЭСМ-6.
 Труды Первой всесоюзной конференции по программированию, Киев. 1968.
- 2. Бабецкий Г.М., Бежанова М.М. и др. АЛЬФА-система автоматизации программирования, под редакцией Ершова. Новосибирск, 1965.
- 3. Волков А.И. Автокод марыем , препринт ОИЯИ, В4-II-46-54, Дубна, 1969.
- 4. Н.Н.Говорун, В.Ю.Веретенов, А.И.Волков, Н.С.Заикин, И.Н.Силин, Р.Н.Федорова, В.П.Шириков. Мониторная система ДУБНА для ЭВМ БЭСМ-6. Труды второй всесоюзной конференции по программированию, Новосибирск, 1970.
- 5. Н.П.Трифонов. Язык символического кодирования SIBESM -6 (Методическое пособие для программирования), МГУ, М, 1967.

Рукопись поступила в издательский отдел 26 октября 1970 года.

СОДЕРЖАНИЕ

Введение	стр 3
I. Обращение к транслятору	3
2. Две формы инструкций	4.
3. Состав транслятора	6
4. Схема трансляции	9
5. Используемые массивы	II.
6. Некоторые детали	14
Приложение I Массив <i>RABINF</i>	15
Приложение 2	
Подпрограмма для работы с лентами символов	16
Литература	19