

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МОЭВМ

ОТЧЕТ
по лабораторной работе №1
по дисциплине «Операционные системы»
Тема: Исследование структур загрузочных модулей

Студент гр. 0381

Шыныбаев А

Преподаватель

Ефремов М. А.

Санкт-Петербург

2022

Цель работы.

Исследование различий в структурах исходных текстов модулей типов .COM и .EXE, структур файлов загрузочных модулей и способов их загрузки в основную память.

Постановка задачи.

Требуется написать текст исходного .COM модуля, который определяет тип PC и версию системы. Ассемблерная программа должна читать содержимое предпоследнего байта ROM BIOS, по таблице, сравнивая коды, определять тип PC и выводить строку с названием модели. Если код не совпадает ни с одним значением, то двоичный код переводиться в символьную строку, содержащую запись шестнадцатеричного числа и выводиться на экран в виде соответствующего сообщения. Затем определяется версия системы. Ассемблерная программа должна по значениям регистров AL и AH сформировать текстовую строку в формате xx.yy, где xx – номер основной версии, а yy - номер модификации в десятичной системе счисления, сформировать строки с серийным номером OEM (Original Equipment Manufacturer) и серийным номером пользователя. Полученные строки выводятся на экран. Далее необходимо отладить полученный исходный модуль и получить «хороший» .COM модуль, а также необходимо построить «плохой» .EXE, полученный из исходного текста для .COM модуля. Затем нужно написать текст «хорошего» .EXE модуля, который выполняет те же функции, что и модуль .COM, далее его построить, отладить и сравнить исходные тексты для .COM и .EXE модулей.

Таблица 1 — Процедуры в программе.

Процедура	Описание
TETR_TO_HEX	Перевод десятичной цифры в код символа
BYTE_TO_HEX	Перевод байта в 16-ной с/с в символьный код
WRD_TO_HEX	Перевод слова в 16-ной с/с в символьный код
BYTE_TO_DEC	Перевод байта в 16-ной с/с в символьный код в 10-ной с/с
model_print	Вывод строки на экран
PC_ver	Определение модели PC
OS_ver	Определение версии OS
OEM_num	Определение OEM
USER_num	Определение серийного номера пользователя

Выполнение работы.

Данные объявленные в программе:

PCm db 'PC',0Dh,0Ah,'\$'

XTm db 'PC/XT',0Dh,0Ah,'\$'

ATm db 'AT',0Dh,0Ah,'\$'

PS2_30m db 'PS2 model 30',0Dh,0Ah,'\$'

PS2_80m db 'PS2 model 80',0Dh,0Ah,'\$'

PS_jrm db 'PCjr',0Dh,0Ah,'\$'

PCconv_m db 'PC Convertible',0Dh,0Ah,'\$'

PCcust_m db ' ', 0Dh, 0Ah, '\$'

USER db ' ' , '\$'

Далее представлены скриншоты полученных модулей.

Рис.1 - хороший СОМ модуль

Рис. 2 - плохой EXE модуль

Рис.3 - хороший EXE модуль

Отличия исходных текстов COM и EXE программ

1. Сколько сегментов должна содержать COM-программа?

- Один, код и данные в COM-модуле располагаются в одном сегменте, а стек генерируется автоматически.

2. EXE-программа?

- EXE-модуль должен содержать сегмент кода и сегмент данных. Остальные сегменты являются опциональными. Если не объявить стек, то будет использоваться DOS-овский.

3. Какие директивы должны быть обязательно в тексте COM-программы?

- 1. `ORG 100h` - Так как адресация начинается с шест. смещения 100 от начала PSP, то в программе после оператора `SEGMENT` кодируется директива `ORG 100H`.
- 2. `ASSUME` - для того, чтобы сегмент данных и сегмент кода указывали на один общий сегмент.

4. Все ли форматы команд можно использовать в COM-программе?

- Нет. Нельзя использовать команды вида `mov register, segment`, т. к. В момент ассемблирования и редактирования связей сегментное значение для сегмента неизвестно. Оно определяется только при загрузке программы. Поскольку файл типа `.COM` не может предоставить загрузчику перечня всех сегментных ссылок (информация для перемещения), то в данном случае программа будет выполняться неправильно.

Отличия форматов файлов .COM и .EXE программ

1. Какова структура файла .COM? С какого адреса располагается код?

Address	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f	Dump
00000000	e9	80	01	50	43	0d	0a	24	50	43	2f	58	54	0d	0a	24	йЪ.РС..\$PC/XT..\$
00000010	41	54	0d	0a	24	50	53	32	20	6d	6f	64	65	6c	20	33	AT..\$PS2 model 3
00000020	30	0d	0a	24	50	53	32	20	6d	6f	64	65	6c	20	38	30	0..\$PS2 model 80
00000030	0d	0a	24	50	d0	a1	6a	72	0d	0a	24	50	43	20	43	6f	..\$PPÿjr..\$PC Co
00000040	6e	76	65	72	74	69	62	6c	65	0d	0a	24	20	20	0d	0a	nvertible..\$..
00000050	24	20	20	2e	20	20	0d	0a	24	20	20	20	0d	0a	24	20	\$. ..\$..\$
00000060	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	
00000070	20	20	20	20	24	24	0f	3c	09	76	02	04	07	04	30	c3	\$\$.<.v....0Г
00000080	51	8a	e0	e8	ef	ff	86	c4	b1	04	d2	e8	e8	e6	ff	59	QЪаипя†Д±.ТиижяУ
00000090	c3	53	8a	fc	e8	e9	ff	88	25	4f	88	05	4f	8a	c7	e8	ГSЪийя€%O€.OЪЗи
000000a0	de	ff	88	25	4f	88	05	5b	c3	51	52	3c	00	74	1f	32	юя€%O€. [ГQR<.t.2
000000b0	e4	33	d2	b9	0a	00	f7	f1	80	c2	30	88	14	4e	33	d2	дЗТ№...чсЪВ0€.NЗТ
000000c0	3d	0a	00	73	f1	3c	00	0c	30	88	04	eb	05	90	0c	30	=...зс<...O€.л...0
000000d0	88	04	5a	59	c3	b4	09	cd	21	c3	b8	00	f0	8e	c0	26	€.ZYГг.Н!Гё.pЪA&
000000e0	a0	fe	ff	3c	ff	74	1f	3c	fe	74	21	3c	fb	74	1d	3c	.юя<ят.<ют!<ыт.<
000000f0	fc	74	1f	3c	fa	74	21	3c	f8	74	23	3c	fd	74	32	3c	ьт.<ът!<шт#<эт2<
00000100	f9	74	34	eb	1f	90	ba	03	01	eb	2f	90	ba	08	01	eb	щт4л...е...л/.е...л
00000110	29	90	ba	10	01	eb	23	90	ba	15	01	eb	1d	90	ba	24) .е...л#.е...л...е\$
00000120	01	eb	17	90	be	4c	01	46	e8	55	ff	ba	4c	01	eb	0a	.л...sL.ФиУяєL.л.
00000130	90	ba	33	01	eb	04	90	ba	3b	01	e8	98	ff	c3	b4	30	.єЗ.л...е; .и.яГг0
00000140	cd	21	50	be	51	01	46	e8	5f	ff	58	83	c6	04	e8	58	Н!РзQ.Фи_яХѳЖ.иХ
00000150	ff	ba	51	01	e8	7e	ff	c3	be	59	01	8a	c7	e8	49	ff	яєQ.и~яГзУ.ЪЗиІя
00000160	ba	59	01	e8	6f	ff	c3	bf	5f	01	83	c7	05	8b	c1	e8	еУ.иояГі_.ѳЗ.<Би
00000170	1f	ff	8a	c3	e8	09	ff	83	ef	02	89	05	ba	5f	01	e8	.яЪГи.яѳп.%.е_.и
00000180	53	ff	c3	e8	54	ff	e8	b5	ff	e8	cc	ff	e8	d8	ff	32	СяГиТяириМяиШя2
00000190	c0	b4	4c	cd	21												ArLH!

Рис.4 - хороший com модуль

- COM-файл состоит из одного сегмента, сегмент стека генерируется автоматически при создании COM - модуля.
- COM-файл ограничен размером одного сегмента и не превышает 64 Кб.
- Программа, записанная в файле типа .COM может сразу выполняться (из-за постоянного смещения).
- Код начинается с адреса 0h, но при загрузке модуля устанавливается смещение в 100h. (в доказательство приведен рис. 4)

2. Какова структура файла плохого EXE? С какого адреса располагается код?

Что располагается с адреса 0?


```

00000300 e9 80 01 50 43 0d 0a 24 50 43 2f 58 54 0d 0a 24 йЪ.PC..$PC/XT..$
00000310 41 54 0d 0a 24 50 53 32 20 6d 6f 64 65 6c 20 33 AT..$PS2 model 3
00000320 30 0d 0a 24 50 53 32 20 6d 6f 64 65 6c 20 38 30 0..$PS2 model 80
00000330 0d 0a 24 50 d0 a1 6a 72 0d 0a 24 50 43 20 43 6f ..$PPŸjr..$PC Co
00000340 6e 76 65 72 74 69 62 6c 65 0d 0a 24 20 20 0d 0a nvertible..$ ..
00000350 24 20 20 2e 20 20 0d 0a 24 20 20 20 0d 0a 24 20 $ . ..$ ..$
00000360 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
00000370 20 20 20 20 24 24 0f 3c 09 76 02 04 07 04 30 c3 $$.<.v....0Г
00000380 51 8a e0 e8 ef ff 86 c4 b1 04 d2 e8 e8 e6 ff 59 QЪаипя†Д±.ТиикяŸ
00000390 c3 53 8a fc e8 e9 ff 88 25 4f 88 05 4f 8a c7 e8 ГSЪыййя€%O€.OЪSi
000003a0 de ff 88 25 4f 88 05 5b c3 51 52 3c 00 74 1f 32 юя€%O€. [ГQR<.t.2
000003b0 e4 33 d2 b9 0a 00 f7 f1 80 c2 30 88 14 4e 33 d2 дзТЙ..чсЪВ0€.N3T
000003c0 3d 0a 00 73 f1 3c 00 0c 30 88 04 eb 05 90 0c 30 =..sc<..0€.л...0
000003d0 88 04 5a 59 c3 b4 09 cd 21 c3 b8 00 f0 8e c0 26 €.ZYГг.Н!Гё.pЪAа
000003e0 a0 fe ff 3c ff 74 1f 3c fe 74 21 3c fb 74 1d 3c .юя<ят.<ѳт!<ѳт.<
000003f0 fc 74 1f 3c fa 74 21 3c f8 74 23 3c fd 74 32 3c ѳт.<ѳт!<ѳт#<ѳт2<
00000400 f9 74 34 eb 1f 90 ba 03 01 eb 2f 90 ba 08 01 eb щт4л..е..л/.е..л
00000410 29 90 ba 10 01 eb 23 90 ba 15 01 eb 1d 90 ba 24 ).е..л#.е..л..е$
00000420 01 eb 17 90 be 4c 01 46 e8 55 ff ba 4c 01 eb 0a .л..sL.ГиUяeL.л.
00000430 90 ba 33 01 eb 04 90 ba 3b 01 e8 98 ff c3 b4 30 .е3.л..е;.и.яГг0
00000440 cd 21 50 be 51 01 46 e8 5f ff 58 83 c6 04 e8 58 Н!PsQ.Ги_яХѳЖ.иX
00000450 ff ba 51 01 e8 7e ff c3 be 59 01 8a c7 e8 49 ff яеQ.и~яГsŸ.ЪSiIя
00000460 ba 59 01 e8 6f ff c3 bf 5f 01 83 c7 05 8b c1 e8 еŸ.иояГг_.ѳЪ.<Би
00000470 1f ff 8a c3 e8 09 ff 83 ef 02 89 05 ba 5f 01 e8 .яЪГи.яѳп.%.е_.и
00000480 53 ff c3 e8 54 ff e8 b5 ff e8 cc ff e8 d8 ff 32 СяГиТяицямМяиШя2
00000490 c0 b4 4c cd 21 AgLH!

```

Рис. 5 - плохой EXE модуль

- У «плохого» EXE файла данные и код располагаются в одном сегменте, однако это не соответствует формату EXE.
- Код начинается с адреса 300h, а с адреса 0h идёт настраивающая таблица (заголовок EXE файла). (Это иллюстрирует Рис. 5)

3. Какова структура хорошего EXE? Чем он отличается от файла плохого EXE?


```

00000300 50 43 0d 0a 24 50 43 2f 58 54 0d 0a 24 41 54 0d PC..$PC/XT..$AT.
00000310 0a 24 50 53 32 20 6d 6f 64 65 6c 20 33 30 0d 0a . $PS2 model 30..
00000320 24 50 53 32 20 6d 6f 64 65 6c 20 38 30 0d 0a 24 $PS2 model 80..$
00000330 50 d0 a1 6a 72 0d 0a 24 50 43 20 43 6f 6e 76 65 PPŸjr..$PC Conve
00000340 72 74 69 62 6c 65 0d 0a 24 20 20 0d 0a 24 20 20 rtible..$ ..$
00000350 2e 20 20 0d 0a 24 20 20 20 0d 0a 24 20 20 20 20 . ..$ ..$
00000360 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
00000370 20 24 00 00 00 00 00 00 00 00 00 00 00 00 00 00 $.....
00000380 24 0f 3c 09 76 02 04 07 04 30 c3 51 8a e0 e8 ef $.<.v....0GQЪаип
00000390 ff 86 c4 b1 04 d2 e8 e8 e6 ff 59 c3 53 8a fc e8 я†Д±.ТиижяYTSЪьи
000003a0 e9 ff 88 25 4f 88 05 4f 8a c7 e8 de ff 88 25 4f йя€%O€.OЪзиЮя€%O
000003b0 88 05 5b c3 51 52 3c 00 74 1f 32 e4 33 d2 b9 0a €. [GQR<.t.2д3TŦ.
000003c0 00 f7 f1 80 c2 30 88 14 4e 33 d2 3d 0a 00 73 f1 .чсЪB0€.N3T=..sc
000003d0 3c 00 0c 30 88 04 eb 05 90 0c 30 88 04 5a 59 c3 <..0€.л...0€.ZYГ
000003e0 b4 09 cd 21 c3 b8 00 f0 8e c0 26 a0 fe ff 3c ff г.Н!Гё.рѦА€.юя<я
000003f0 74 1f 3c fe 74 21 3c fb 74 1d 3c fc 74 1f 3c fa т.<ѣт!<ѣт.<ѣт.<ѣ
00000400 74 21 3c f8 74 23 3c fd 74 32 3c f9 74 34 eb 1f т!<ѣт#<ѣт2<ѣт4л.
00000410 90 ba 00 00 eb 2f 90 ba 05 00 eb 29 90 ba 0d 00 .е..л/.е..л).е..
00000420 eb 23 90 ba 12 00 eb 1d 90 ba 21 00 eb 17 90 be л#.е..л..е!.л..s
00000430 49 00 46 e8 55 ff ba 49 00 eb 0a 90 ba 30 00 eb I.ФиУяеI.л..е0.л
00000440 04 90 ba 38 00 e8 98 ff c3 b4 30 cd 21 50 be 4e ..е8.и.яГг0Н!PsN
00000450 00 46 e8 5f ff 58 83 c6 04 e8 58 ff ba 4e 00 e8 .Фи_яХѦЖ.иХяеN.и
00000460 7e ff c3 be 56 00 8a c7 e8 49 ff ba 56 00 e8 6f ~яГsV.ЪзиIяеV.ио
00000470 ff c3 bf 5c 00 83 c7 05 8b c1 e8 1f ff 8a c3 e8 яГi\.\Ѧ9.<Би.яЪГи
00000480 09 ff 83 ef 02 89 05 ba 5c 00 e8 53 ff c3 2b c0 .яѦп.%.е\..иSяГ+A
00000490 50 b8 10 00 8e d8 e8 4c ff e8 ad ff e8 c4 ff e8 Рё..ѦѦиLяи-яиДяи
000004a0 d0 ff 32 c0 b4 4c cd 21 Ря2ArLH!

```

Рис. 6 - хороший EXE модуль

- В EXE-модуле код и данные являются отдельными сегментами, также присутствует таблица связей, заголовок, отвечающий за настройку адресов.
- В «хорошем» EXE-модуле происходит разделение сегментов (кода и данных), необходимое для правильного форматирования, а в «плохом» содержится лишь один сегмент, объединяющий код и данные. «Плохой» EXE начинает код с 300h, так как он получается из COM модуля, в котором изначально сегмент кода смещён на 100h, Но, так как, происходит создание EXE-модуля, добавляется еще и сдвиг PSP (200h). В «хорошем» EXE присутствует только смещение для PSP модуля, поэтому код начинается с 200h.

- В данном случае смещение кода 300h так как выделяется память под стек (в размере 100h), память под стек находится между PSP и кодом. (Как показано на рис. 6)
4. Как определяется стек? Какую область памяти он занимает? Какие адреса?
- Стек находится между PSP и данными и занимает с 100h до 300h

Загрузка COM модуля в основную память

1. Какой формат загрузки модуля COM? С какого адреса располагается код?
- Определяется сегментный адрес участка Основной Памяти, у которого достаточно места для загрузки программы, образ COM-файла считывается с диска и помещается в память, начиная с адреса PSP 100h. После загрузки двоичного образа COM-модуля сегментные регистры CS, DS, ES и SS указывают на PSP(в данном случае сегментные регистры указывают на 48DD), SP указывает на конец сегмента PSP (FFFE), слово 00H помещено в стек, IP содержит 100H. (Это можно увидеть на Рис. 7)

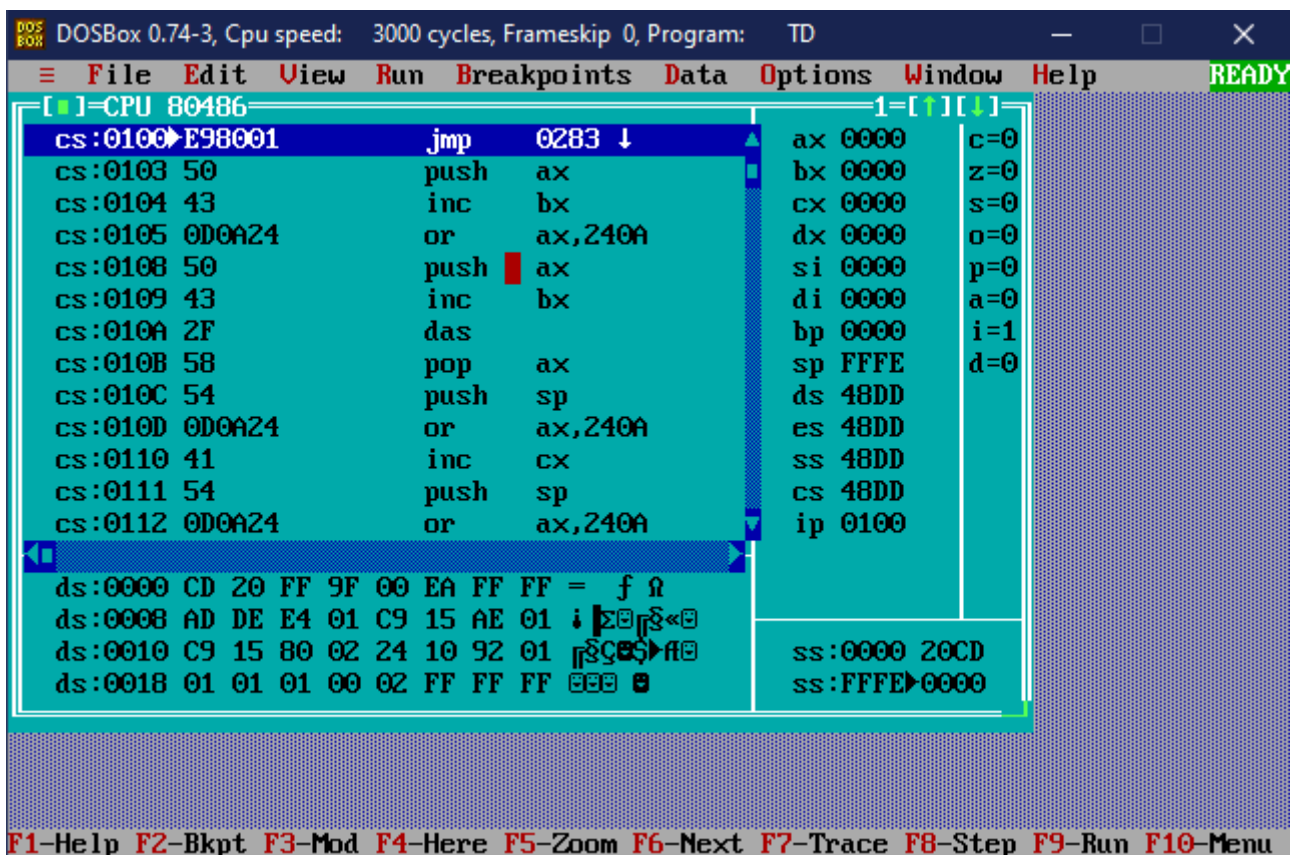


Рис. 7 - отладчик для хорошего COM-модуля

2. Что располагается с адреса 0?

- Программный сегмент PSP, размером 256 байт (100h), зарезервируемый операционной системой.

3. Какие значения имеют сегментные регистры? На какие области памяти они указывают?

- Сегментные регистры CS, DS, ES и SS указывают на PSP и имеют значения 48DD. (Это можно увидеть на Рис. 7)

4. Как определяется стек? Какую область памяти он занимает? Какие адреса?

- Стек генерируется автоматически при создании COM-программы. SS – на начало (0h), регистр SP указывает на конец стека (FFFEh), Адреса стека расположены в диапазоне 0h – FFFEh (FFFEh, – последний адрес, кратный двум). (Это можно увидеть на Рис. 7)

Загрузка «хорошего» EXE модуля в основную память

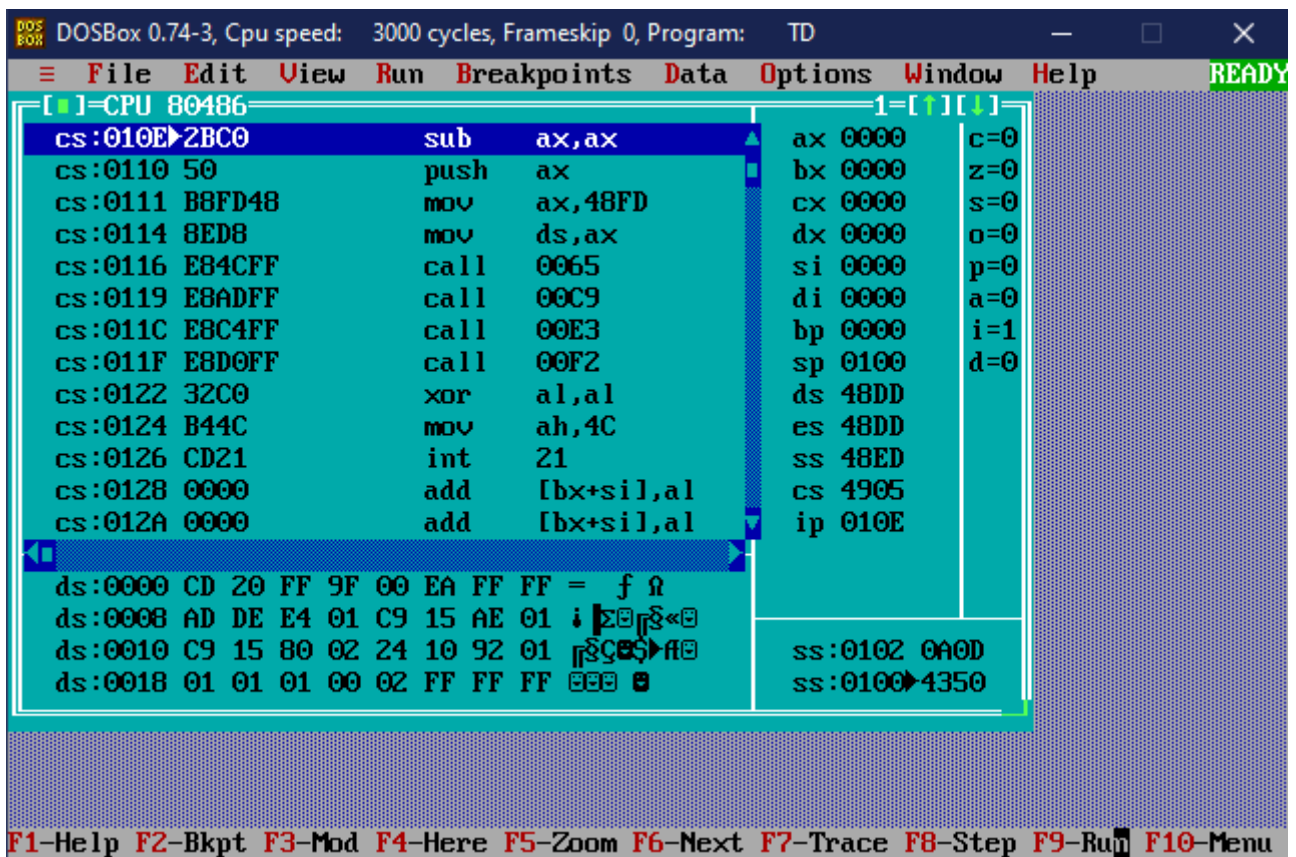


Рис. 8 - Отладчик "хорошего" EXE-модуля

1. Как загружается «хороший» .EXE? Какие значения имеют сегментные регистры?

- EXE-файл загружается, начиная с адреса PSP:0100h. В процессе загрузки считывается информация заголовка (PSP) EXE в начале файла и выполняется перемещение адресов сегментов, то есть DS и ES устанавливаются на начало сегмента PSP(DS=ES=48DD), SS(SS=48ED) – на начало сегмента стека, CS(CS=4905) – на начало сегмента команд. В IP загружается смещение точки входа в программу, которая берётся из метки после директивы END. (Это иллюстрирует рисунок 8)

2. На что указывают регистры DS и ES?

- Регистры DS и ES указывают на начало PSP.

3. Как определяется стек?

- Стек определяется с помощью Stack Segment, после которой задается размер стека. При исполнении регистр SS указывает на начало сегмента стека, а SP на конца стека(его смещение).

4. Как определяется точка входа?

- Точка входа определяется при помощи директивы END.

Выводы.

Были написаны COM и EXE модули, на основе которых производилось сравнение данных форматов. Также были выявлены недостатки и преимущества каждого из них.