

Operating Systems Concepts

Chapter 5: **Programs as Data**

- Programs that manipulate programs
- Assemblers, compilers
- Linking, name binding
- Relocation

Olav Beckmann

Huxley 449

<http://www.doc.ic.ac.uk/~ob3>

Acknowledgements: There are lots. See end of Chapter 1.

Home Page for the course:

<http://www.doc.ic.ac.uk/~ob3/Teaching/OperatingSystemsConcepts/>

This is only up-to-date after I have issued printed version of the notes, tutorials, solutions etc.

Chapter 5: Programs as Data

The purpose of this chapter:

- Assemblers and the assembly language representation of machine code
- Compiling high-level language to machine code
- How names (such as variable names) are bound to concrete values
- Linking separate program components; name binding
- What has to happen when a program is loaded
- Textbook: Nutt page 120-121 and 419-429

Programming the NARC

- You program a computer by putting just the right numbers into its memory
- This means that early programmers had to be incredibly clever
- Programming has been de-skilled
- Some simple tools mean that anyone can do it 😊
- Key tools:
 - Assembler
 - Compiler
 - Linker

Compiler, Assembler, Linker...

- The funny thing is that these three words all seem to mean the same thing
- In computing, all definitions are flexible, but they have come to refer to the following structure:

- **Assembler**: translates human-readable versions of machine instructions into the machine encoding, ready for direct interpretation by the processor
- **Compiler**: translates a high-level language (C, C++, etc) into machine instructions
- **Linker**: combines chunks of machine instructions (e.g. separately purchased software) together

Assembler

Assembler Source

A	.word	23
B	.word	45
C	.word	0
main	loadm	A
	jmpz	end
	loadm	C
	addm	B
	storem	C
	loadm	A
	subc	1
	storem	A
	jmp	main
end		

Object Code

23		0
45		1
0		2
2	0	3
9	?	4
2	2	5
5	1	6
3	2	7
2	0	8
6	1	9
3	0	10
8	3	11
		12

Textual representation, one line per instruction

Instructions and data, encoded in binary, ready to be operated upon directly by the processor

Assembler Source

A	.word	23
B	.word	45
C	.word	0
main	loadm	A
	jmpz	end
	loadm	C
	addm	B
	storem	C
	loadm	A
	subc	1
	storem	A
	jmp	main
end		

Object Code

23		0
45		1
0		2
2	0	3
9	12	4
2	2	5
5	1	6
3	2	7
2	0	8
6	1	9
3	0	10
8	3	11
		12

Labels in the assembler source - 'A', 'B', 'C', 'main', 'end' - represent numbers which have to refer to the right address wherever it might turn out to be

As each line of the assembler source is translated, it is assigned an address - so we know the numeric value of 'A', 'B', 'C' - but not 'end'

Two Pass Assembler

Pass 1 - build up Symbol Table

A	.word	23
B	.word	45
C	.word	0
main	loadm	A
	jmpz	end
	loadm	C
	addm	B
	storem	C
	loadm	A
	subc	1
	storem	A
	jmp	main
end		

Symbol Table	
<i>Label</i>	<i>Value</i>
A	0
B	1
C	2
main	3
end	12

Two Pass Assembler

Pass 2

- output Object Code

A	.word	23
B	.word	45
C	.word	0
main	loadm	A
	jmpz	end
	loadm	C
	addm	B
	storem	C
	loadm	A
	subc	1
	storem	A
	jmp	main
end		

Symbol Table	
<i>Label</i>	<i>Value</i>
A	0
B	1
C	2
main	3
end	12

23	0	
45	1	
0	2	
2	0	3
9	12	4
2	2	5
5	1	6
3	2	7
2	0	8
6	1	9
3	0	10
8	3	11
		12

Assembler - Summary

- Assembler input language: textual representation of each machine instruction, one line per instruction
- Assembler language includes “directives” to tell assembler to assign symbolic names (“labels”)
- Also directives to name and set aside working storage (variables)
- Assembler typically operates in two passes:
 - **pass 1**: calculate space occupied, build symbol table
 - **pass 2**: reprocess source using symbol table to fill in symbol values

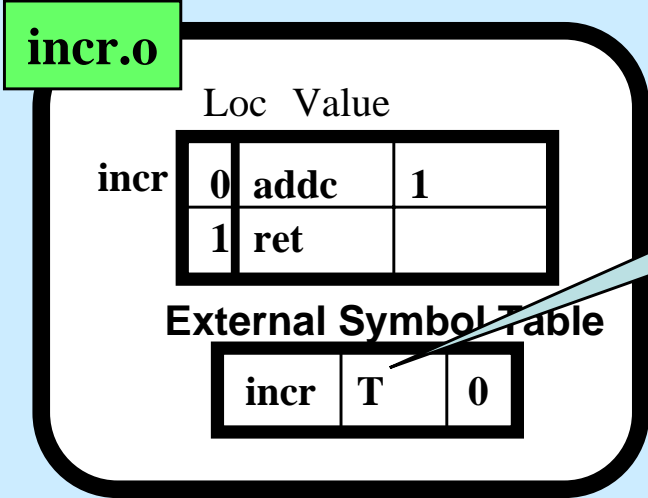
Binding: resolving names

- Two pass assembler:
 - pass 1: calculate space occupied, build symbol table
 - pass 2: reprocess source using symbol table to fill in symbol values
- Two functions:
 - **translate human-readable representation to machine encoding**
 - **resolve references to names**
- **This issue of naming and “binding” of names to values is a recurring theme in operating systems**

The Job of the Linker

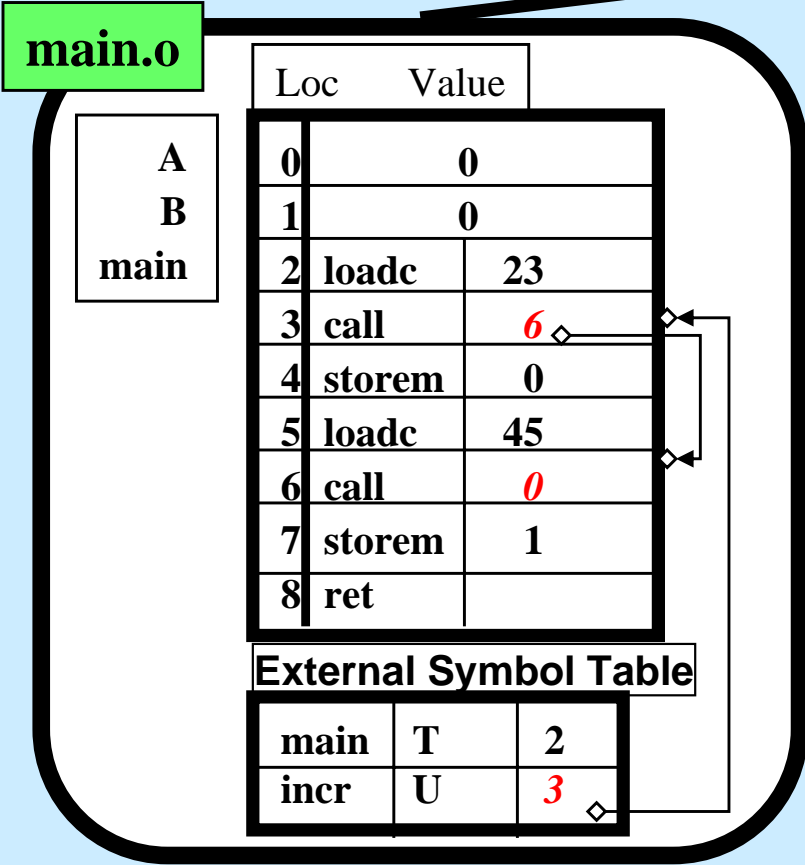
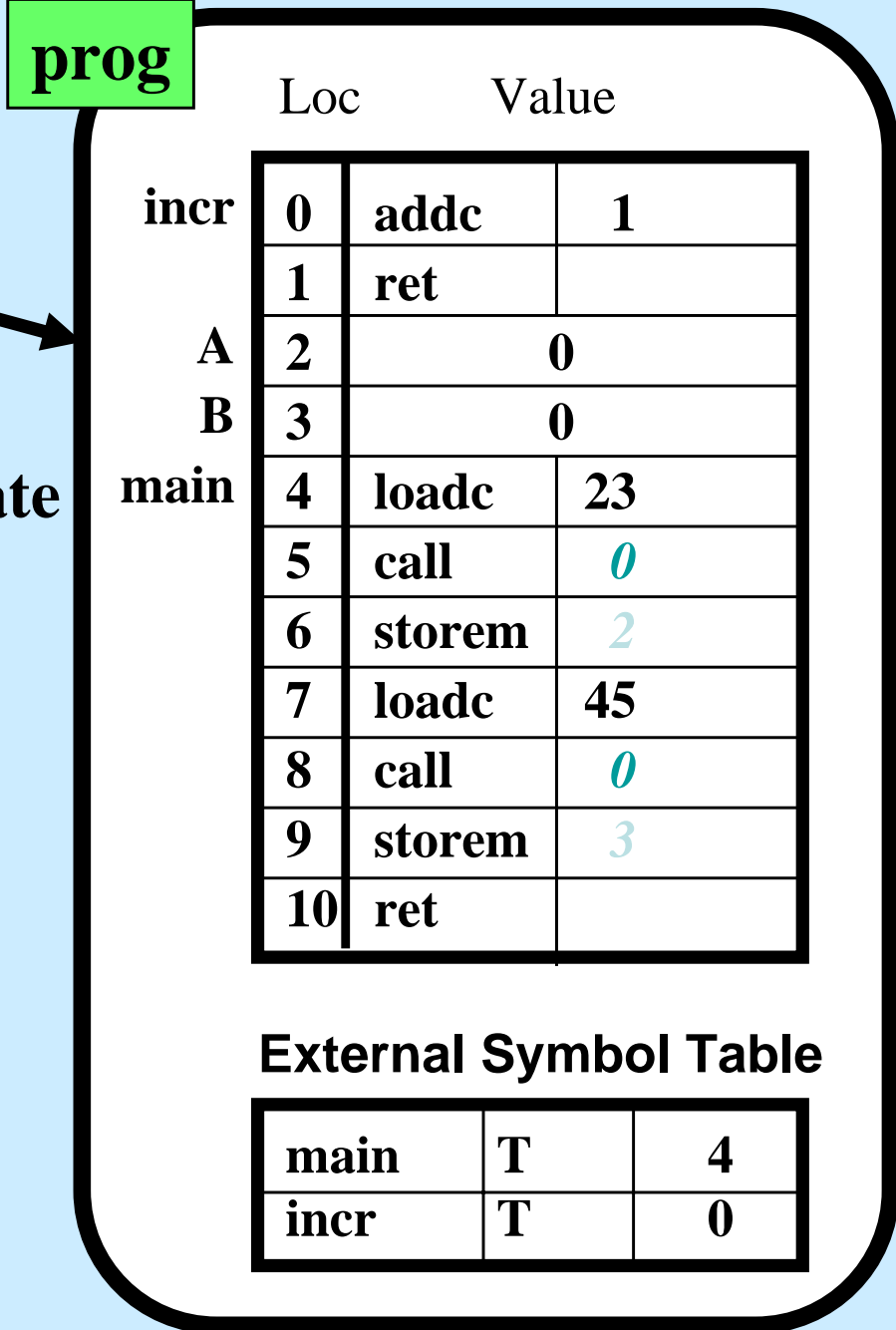
- Suppose we want our new program to use the functionality provided by another program
- E.g. newly-purchased word processor needs to access software driver for new keyboard
- Simpler working example: 'main' program uses separately-provided 'incr' procedure (next slide)
- Two issues:
 - **relocation**: concatenate the binary code - and adjust symbol references according to new addresses
 - **name binding**: resolve names used but undefined in 'main' with names defined in 'incr'

The Linker



T means "Text", referring to program text, rather than data.

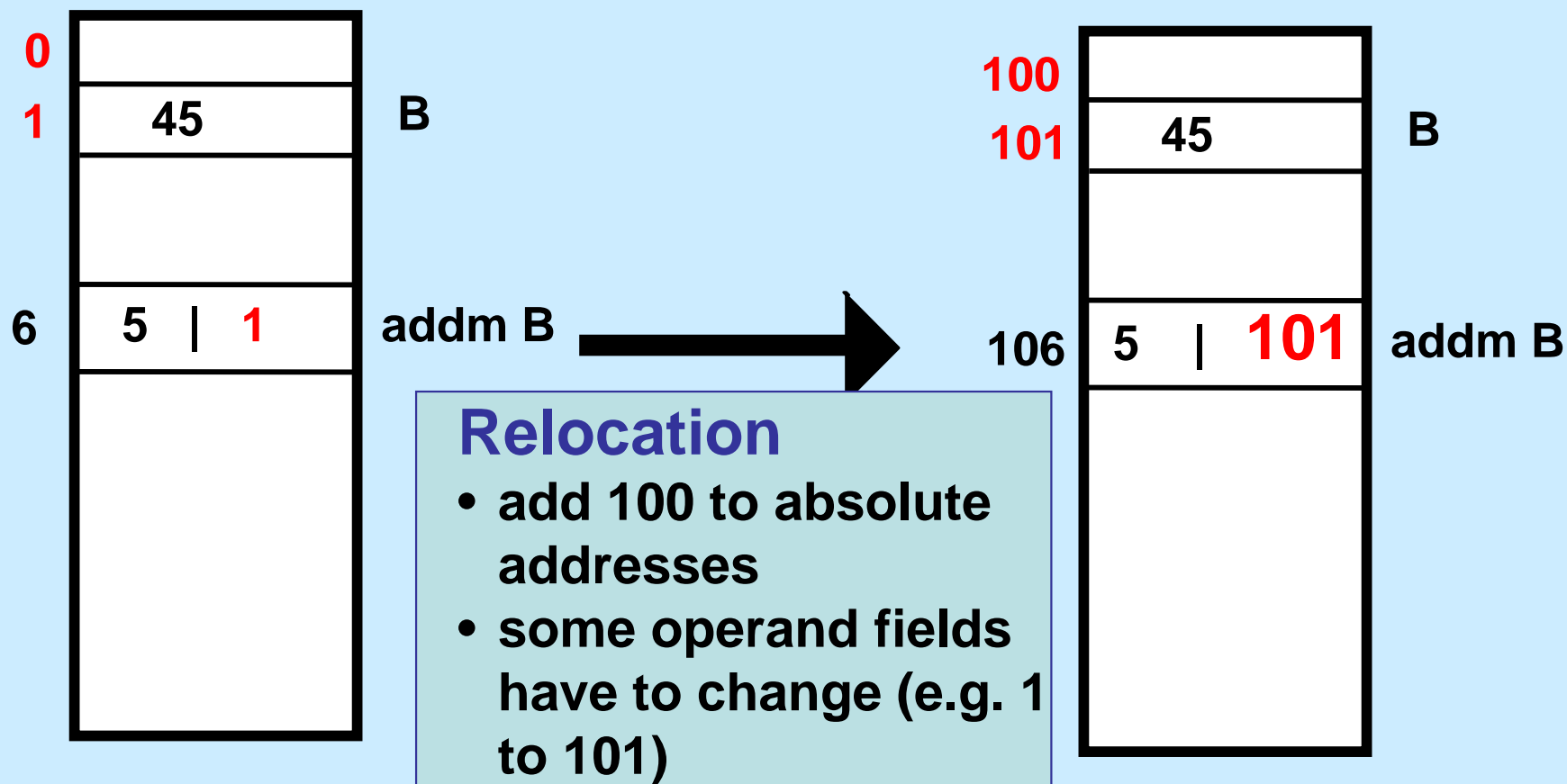
Linker



- 1: Concatenate
- 2: Relocate
- 3: Resolve names

Suppose a program has been assembled starting at address 0.
Suppose it is now loaded into store starting at address 100.

Need to *relocate* absolute address operands



Note: NARC uses only absolute addressing
Relative addresses (e.g. PC relative jumps) do not have to be relocated.

Which locations will need to be adjusted?

Object code file must record this information.

Object Code File

			Loc	Value	Relocate?	
A	.word	23	0	23	n	A
B	.word	45	1	45	n	B
C	.word	0	2	0	n	C
main	loadm	A	3	2	0	y
	jmpz	end	4	9	12	y
	loadm	C	5	2	2	y
	addm	B	6	5	1	y
	storem	C	7	3	2	y
	loadm	A	8	2	0	y
	subc	1	9	6	1	n
	storem	A	10	3	0	y
	jmp	main	11	8	3	y
end			12			

Name Binding: External Symbol Table

incr.o

	Loc	Value	Rel?
incr	0	addc	1
	1	ret	n

External Symbol Table

incr	T	0
------	---	---

T = Code Label
D = Data Label
U = Undefined

value

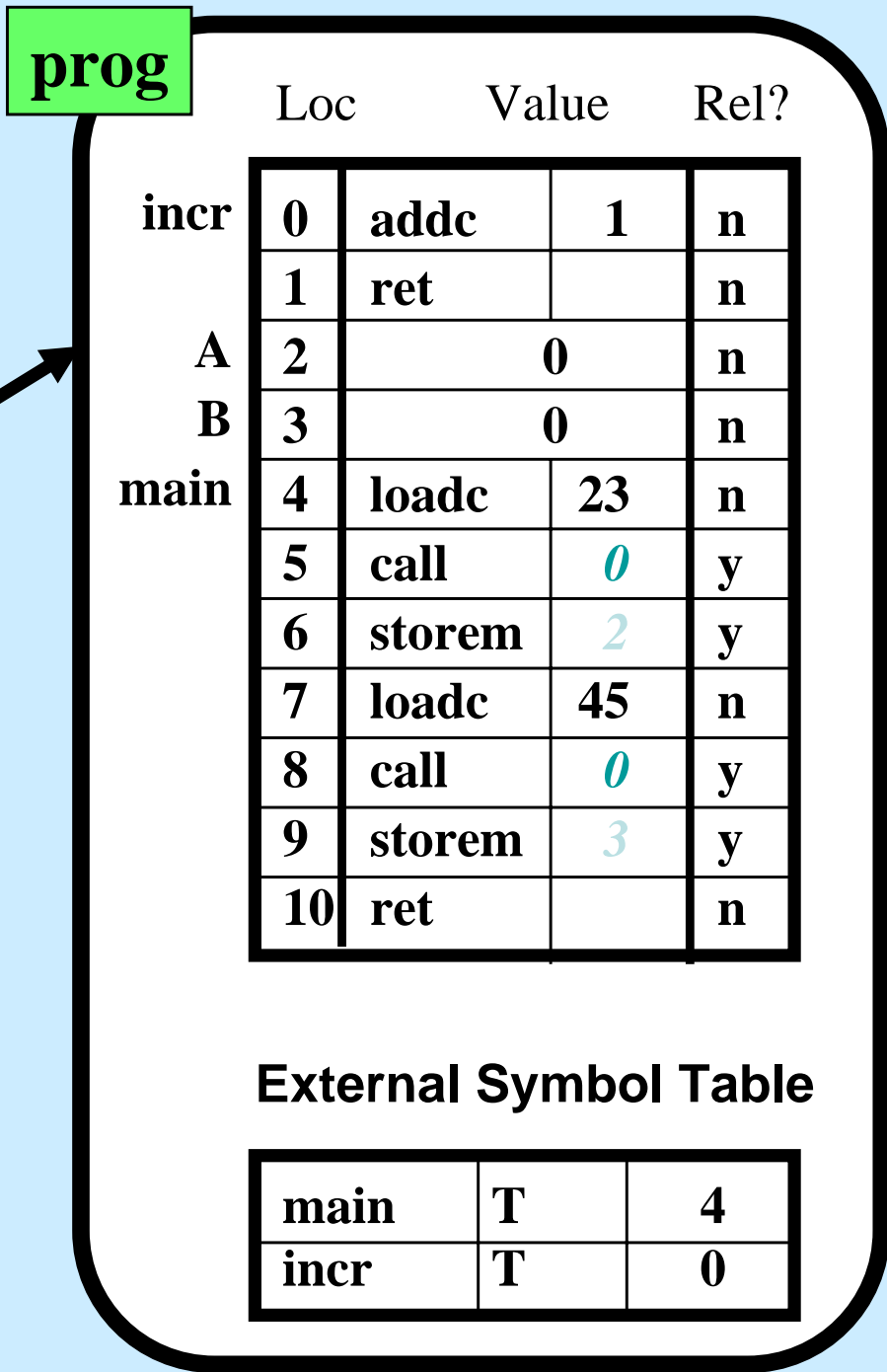
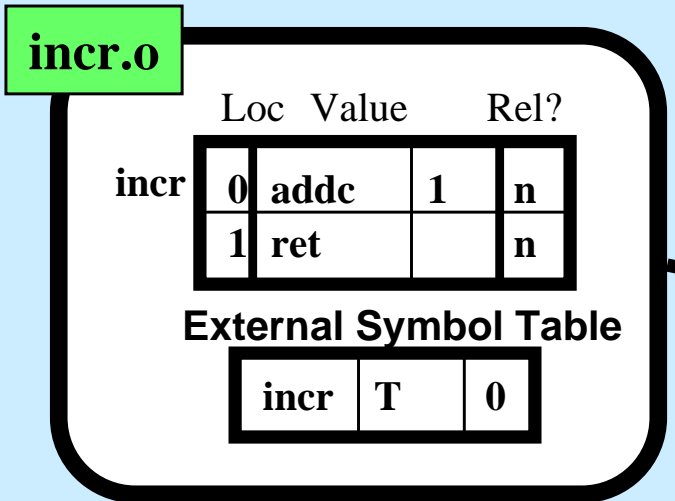
main.o

	Loc	Value	Rel?
A	0	0	n
B	1	0	n
main	2	loadc	23
	3	call	6
	4	storem	0
	5	loadc	45
	6	call	0
	7	storem	1
	8	ret	n

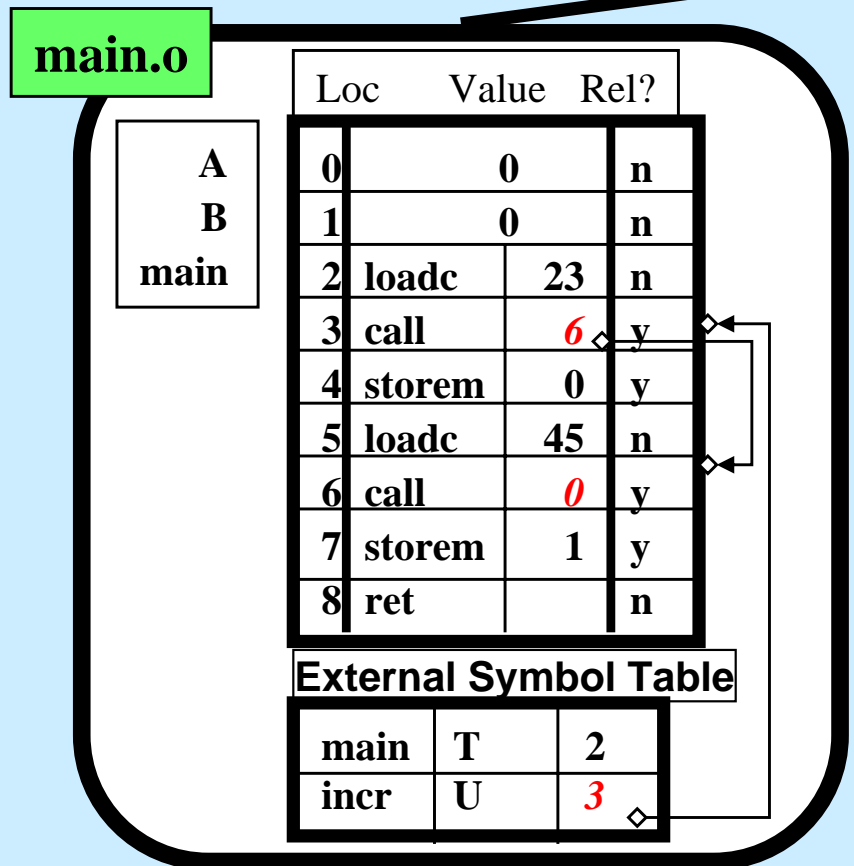
External Symbol Table

main	T	2
incr	U	3

The symbol table entry for an undefined label points to a linked list of entries which use that label



Linker



File Format for Machine Code

- Programs and program components are stored in files
- We need a general-purpose “object code file” format which can represent them
- Need to be able to include:

machine code instructions
values for initialised data
details of uninitialised data space required
names defined
names used but undefined
relocation information
entry point

- An incomplete program doesn't have an entry point
- A complete program has no names used but undefined
- May include further information to aid in debugging

Example - COFF

- Common Object File Format (COFF)
- Used in some Unix systems, basis for Windows Portable Executable format (http://msdn.microsoft.com/library/specs/msdn_pecoff.htm), widely used for embedded systems
- Linux uses ELF (Executable and Linking Format) instead
- To decode an ELF file try using the 'objdump' or 'readelf' command (eg "readelf -a /bin/echo")

Object file format - variations

- When all object files which form a program have been linked, all external references will be resolved and the global symbol table can be discarded.
- However, relocation information must be kept if the program is to be loaded at a different address to that for which it was linked.
- If relocation information is discarded (i.e. loader does not relocate) the program object file is an exact binary image of its representation in main store
- The object file usually records the program entry point for use by the OS when it starts execution of the program - in C/C++ this is the “main” function

Loading a program

- A key function of an OS is to load a user's program and run it
- The program is delivered as an object file, e.g. in COFF format
- Where should the newly-loaded program be put - at which range of memory addresses?

Program Loading - Memory Management

When machine is turned on, PC=0 so it starts execution here

Memory address: 0

Operating system's instructions and data structures

Address of first free word:

Free

Memory address: max

Program Loading - Memory Management

When machine is turned on, PC=0 so it starts execution here

Memory address: 0

Operating system's instructions and data structures

Start address of game:

DeathCopter video game

Once game has been loaded, OS jumps to its start address

Address of first free word:

Free

Memory address: max

Program Loading - Memory Management

When machine is turned on, PC=0 so it starts execution here

Memory address: 0

Operating system's instructions and data structures

Once game has been loaded, OS jumps to its start address

Start address of game:

DeathCopter video game

Start address of player:

MP3 player

What if the MP3 player had been loaded first?

Address of first free word:
Memory address: max

Free

Program Loading - Memory Management

If the order of program loading had been different, the start addresses would be different

Memory address: 0

Operating system's instructions and data structures

Start address of game:

MP3 player

Start address of player:

DeathCopter video game

Address of first free word:
Memory address: max

Free

Relocating loader

prog

memory

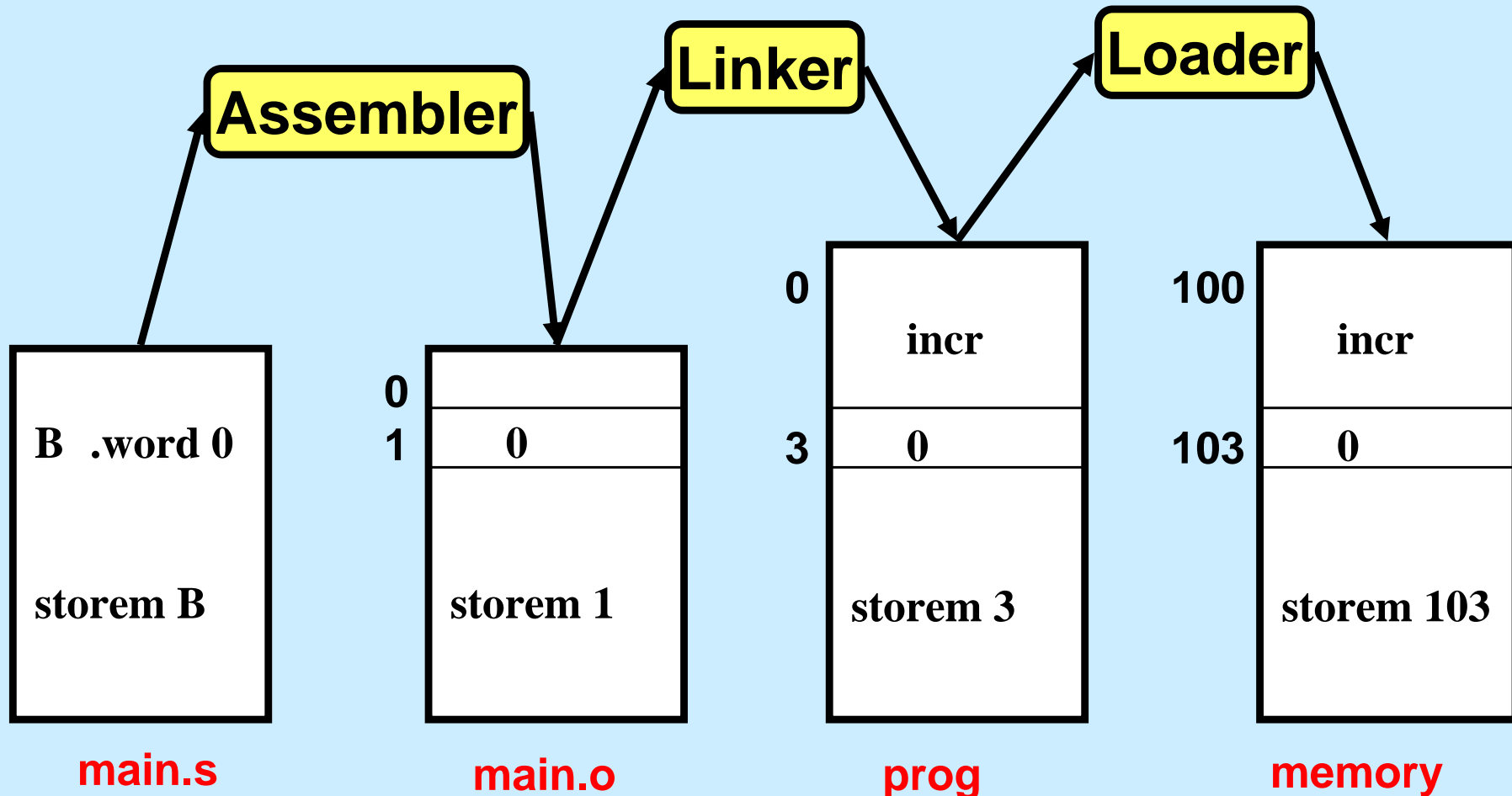
	Loc	Value	Rel?	
incr	0	addc	1	n
	1	ret		n
A	2	0		n
B	3	0		n
main	4	loadc	23	n
	5	call	0	y
	6	storem	2	y
	7	loadc	45	n
	8	call	0	y
	9	storem	3	y
	10	ret		n

entry point = 4

0	...	
1	...	
:	:	
100	addc	1
101	ret	
102		0
103		0
104	loadc	23
105	call	100
106	storem	102
107	loadc	45
108	call	100
109	storem	103
110	ret	
:	:	

Note two stages of relocation:

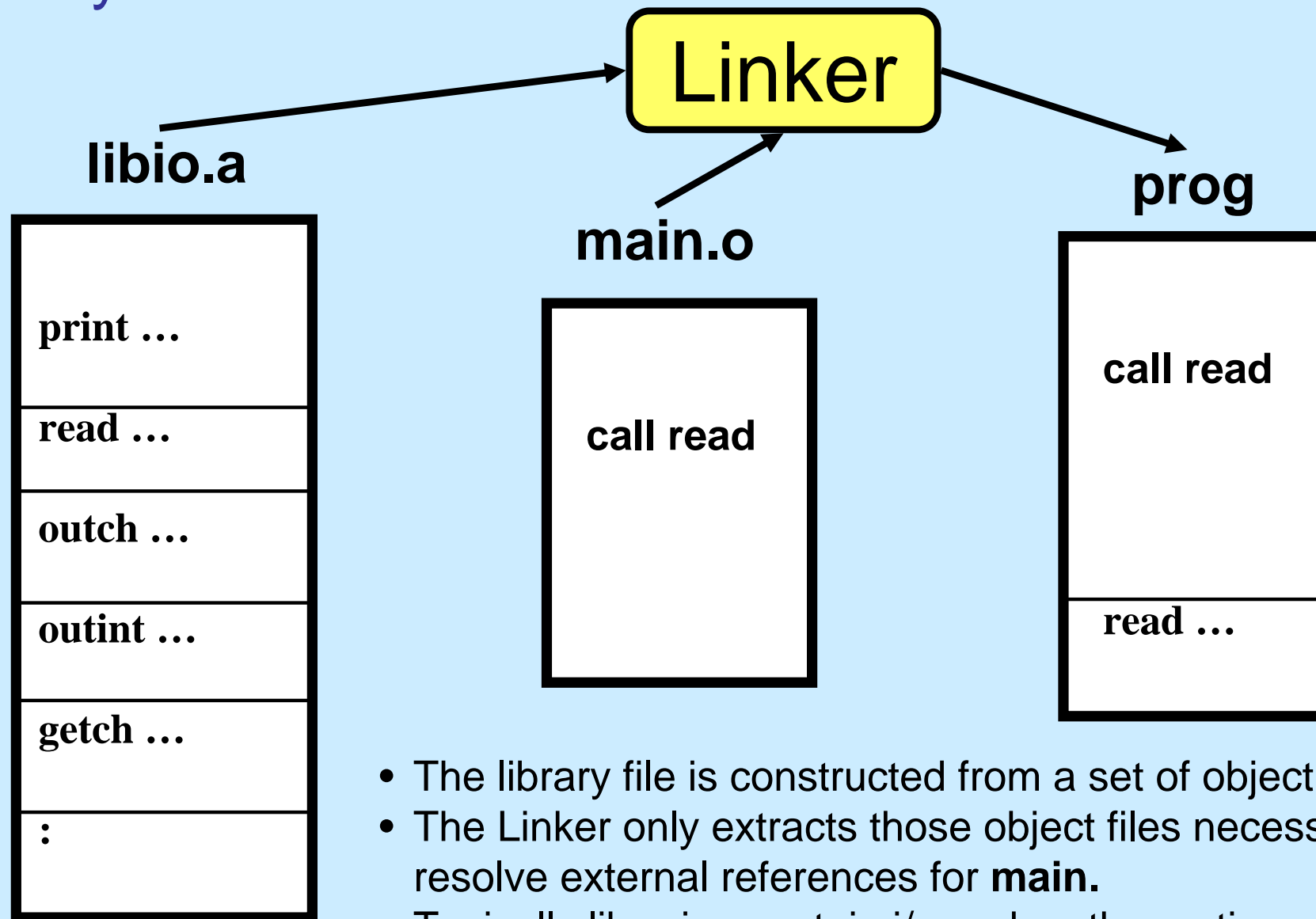
- at link time and load time



Linking, Loading and Relocation - Summary

- Assembler translates human-readable representation of instructions into object file:
 - object file includes list of external names defined, and names used but not defined
 - object file also includes relocation information
- Linker combines object files:
 - concatenate, relocate so internal name references OK
 - resolve name binding
- Loader finds sufficiently-large free memory region, interprets object file format, loads object file instructions and initialised data into memory
 - relocates so that internal name references are right

Library files - constructed by Archiver system utility



- The library file is constructed from a set of object files.
- The Linker only extracts those object files necessary to resolve external references for **main**.
- Typically libraries contain i/o and maths routines.

Using Textbooks

- Nutt and Stallings cover assemblers, compilers and linkers only very briefly
- Eg Nutt page 120-121 and 419-429
- Terminology:
 - Linker = linkage editor
 - External symbol table = external reference table+external definition table
 - In unix/linux, the assembler is “as” and the linker is “ld”

In Real Life...

- This chapter has presented a simplified view
- With **dynamic linking** (Windows DLLs, Linux shared libraries) an object file is loaded during the program's execution
 - Relies on position-independent code
 - Name binding via a table which maps each external reference to its run-time address
- With **just-in-time compilation** (eg Java JIT) object files are represented as machine-independent “bytecode”, which is translated to machine code as it is loaded

In real life... relocation by address translation

- Relocation is not needed if code is “position-independent”
- Relocation is also not needed if the processor has an *address translation* mechanism
- We will cover this in detail later in this course
- The basic idea is that there is a hardware lookup table that intercepts and translates addresses issued by the processor
- This has to be inactive when executing operating system code - which has to be able to set it up