

# CPT210 - Microprocessor

Lecture 6 – ARM Functions and Stacks-Part 1



### Overview

- Lecture 5 Review
  - ARM Assembler
  - Instruction Encoding
  - Rotation

- Lecture 6
  - LDR/STR
    - Notation
    - Instruction encoding
    - Offset
  - ADR
    - Definition
    - Syntax
    - How it works
  - LDR Pseudo-instruction

**Before End: Final Exam & Assessment Explaination** 



- Assemblers
- > Take-away information

Assembler turns readable instructions into real CPU actions.



- Assemblers
- > Take-away information



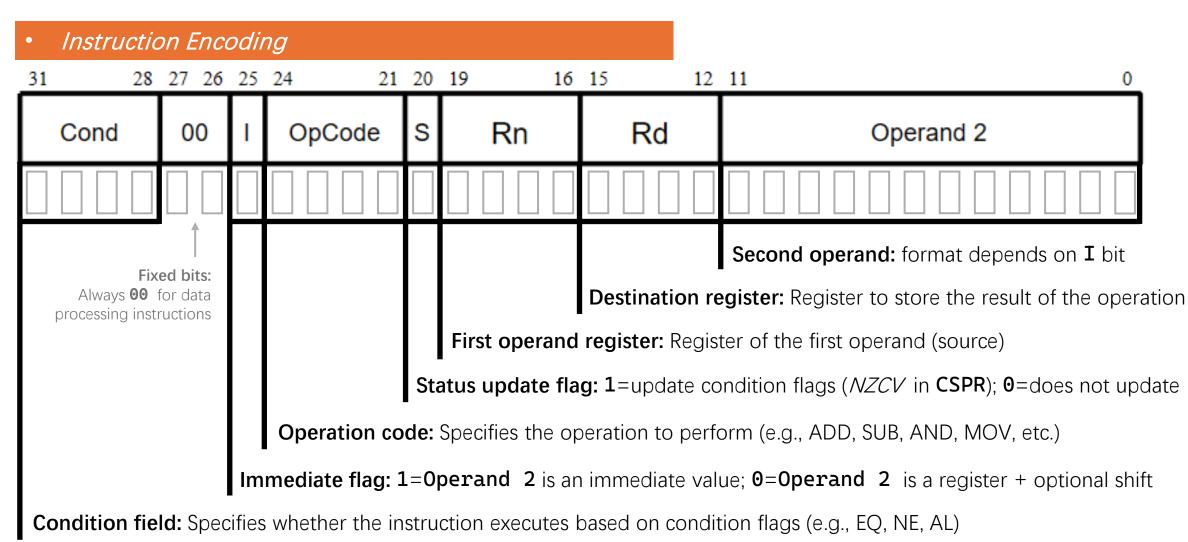
#### Assembler:

- → Translates human-readable assembly code directly into machine code.
- → Used when precise control of hardware and performance is critical.

#### Compiler:

- → Translates **high-level code** into assembly or machine code.
- → Optimizes code for readability, performance, and maintainability.

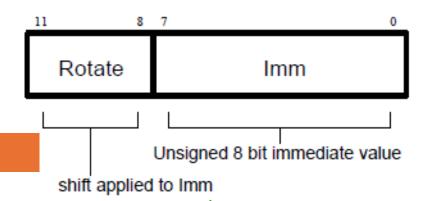






31

### Review of the Lecture 5





Cond 00 I OpCode S

01 = logical right

11 = rotate right

**Shift amount** 

10 = arithmetic right

5 bit unsigned integer

28 27 26 25 24

Rn

16 15

Rd

12 11

11

Operand 2

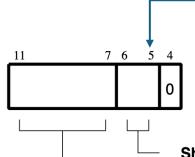
Shift

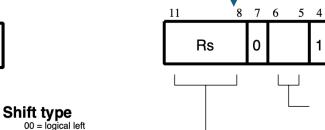
shift applied to Rm

4 3

Rm

2nd operand register





21 20 19

Shift type

00 = logical left
01 = logical right
10 = arithmetic right
11 = rotate right

Shift register

Shift amount specified in bottom byte of Rs



#### Rotation and Shift

#### > Take-away information

Туре	Name	What it does	Example
LSL	Logical Shift Left	Shifts bits left, fills 0s	0010 << 1 = 0100
LSR	Logical Shift Right	Shifts bits right, fills 0s	1000 >> 2 = 0010
ASR	Arithmetic Shift Right	Shifts right, fills with sign bit	1000 >> 2 = 1110 (if signed)
ROR	Rotate Right	Rotates bits around to front	1001 ROR 1 = 1100

**SET OF SET OF S** 



- LDR/STR
- ➤ What are LDR and STR?

Instrcution	Meaning	Basic Function	Syntax
LDR	Load Register	Load a value from memory into a register	LDR Rt, [Rn, #offset]
STR	Store Register	Store the value from a register into memory	STR Rt, [Rn, #offset]



- LDR/STR

ARM follows a **Load/Store architecture**, which means:

- All computations must happen in registers,
- Data in memory cannot be used directly in arithmetic or logic operations.





- LDR/STR
- How to understand LDR and STR?

```
Rt: #offset:

The target register to read from or write to

LDR Rt, [Rn, #offset] Load memory → Rt

STR Rt, [Rn, #offset] Store Rt → memory
```

Rn:

The base register (usually points to the starting memory address)



- LDR/STR
- > Examples

#### Meaning:

Load a 4-byte word from the memory address at R1+4 and store the value into register R0

#### |f R1 = 0x1000|

- Take the address stored in R1, which is 0x1000
- Add #4 (which is 4 bytes) → result: 0x1004
- Read a word starting from address
   0x1004
- Store the loaded word value into register R0

#### • LDR/STR

#### > Try it yourselves

$$R1 = 0 \times 1008$$
,  $R0 = ?$ 

LDR **R0**, [**R1**]

LDR R0, [R1, #8]

LDR R0, [R1, #-8]

Address	Stored Value
0×1000	0xAABBCCDD
0x1008	0x11223344
0×1010	0xDEADBEEF



- LDR/STR
- > Examples

STR R2, [R3]

Meaning:

Store the value in register R2 into the memory address pointed to by R3

one address, one byte

lf

R2 = 0xCAFEBABE

R3 = 0x2000

- Value in register R2 is 0xCAFEBABE
- The memory address pointed to by R3 is 0x2000
- Store the value 0xCAFEBABE at 0x2000

Address	Byte (little-endian)
0×2000	0xBE
0x2001	0xBA
0x2002	0xFE
0x2003	0xCA

ARM uses little-endian format by default, so the lowest byte goes to the lowest address.

#### • LDR/STR

#### > Try it yourselves

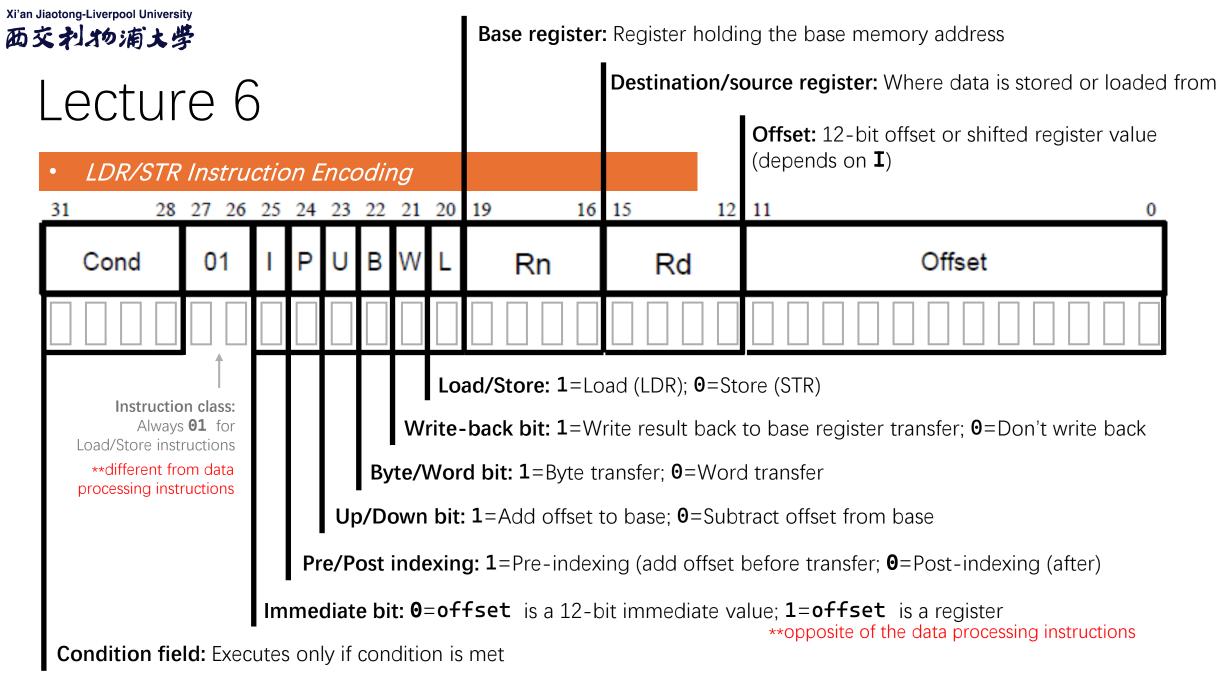
$$R3 = 0 \times 1008$$
,  $R2 = 0 \times CAFEBABE$ 

STR R2, [R3]

STR R2, [R3, #8]

STR R2, [R3, #-8]

Address	Original Stored Value	After overwritting
0x1008	0xAABBCCDD	0xCAFEBABE
0x1010	0x11223344	0xCAFEBABE
0x1000	0xDEADBEEF	0xCAFEBABE



#### • LDR/STR nstruction Encoding

#### Pre/Post indexing (P)

P = 1: pre-indexing, apply the **offset before** the memory access

```
LDR R0, [R1, #4] Pre-index, no write-back LDR R0, [R1, #4]! Pre-index + write-back
```

The memory address = R1 + 4

If you use !, the result (new memory address) is written back to R1

#### > Example

```
R1 = 0 \times 1000
LDR R0, [R1, #8]!
```

- Effective address =  $0 \times 1000 + 8 = 0 \times 1008$
- Load memory at 0x1008 into R0
- Then update R1 =0x1008 (because of!)

- LDR/STR Instruction Encoding
- Pre/Post indexing (P)

P = 0: post-indexing, apply the **offset after** the memory access

LDR R0, [R1], #4 Post-index (write-back implied)

First access memory at R1
Then update R1 = R1 + 4

> Example

 $R1 = 0 \times 1000$ LDR R0, [R1], #8

- Load memort from address
   0x1000 into R0
- Then update R1 = 0x1008 (write-back implied)



#### • LDR/STR Instruction Encoding

#### Pre/Post indexing (P)

Code	Result
LDR R0, [R1, #4]	R0 ← [R1+4], R1 unchanged
LDR R0, [R1, #4]!	$R0 \leftarrow [R1+4], R1 \leftarrow R1+4$
LDR R0, [R1], #4	$R0 \leftarrow [R1], R1 \leftarrow R1+4$

Code	Result
STR R0, [R1, #4]	Store R0 to [R1+4], R1 remains unchanged
STR R0, [R1, #4]!	Store R0 to [R1+4], then update R1 $\leftarrow$ R1 + 4
STR R0, [R1], #4	Store R0 to [R1], then update R1 ← R1 + 4



- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

- What's a "Word" in ARM?
- ARM that we teach is a 32-bit architecture, so a word = 4 bytes = 32 bits
- Word access is aligned to addresses divisible by 4 (e.g., 0x1000, 0x1004...)



- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

Instruction	Transfer Type	Size	Description
LDR	Word (default)	4 bytes	Load 32-bit word from memory
STR	Word (default)	4 bytes	Store 32-bit word to memory
LDR <b>B</b>	Byte	1 byte	Load 8-bit byte from memory
STR <b>B</b>	Byte	1 byte	Store 8-bit byte to memory

◆ The "B" suffix = Byte access

- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

B = 0: Transfer a word (32 bits)

> Example

LDR R0, [R1]

- Loads all 4 bytes starting from 0x1000
- In little-endian: 0xDDCCBBAA → stored in R0

```
R1 = 0x1000

Memory[0x1000] = 0xAA

Memory[0x1001] = 0xBB

Memory[0x1002] = 0xCC

Memory[0x1003] = 0xDD
```

- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

B = 0: Transfer a word (32 bits)

> Example

STR R0, [R1]

- $R0 = 0 \times AABBCCDD$
- R1 = 0x2000

- Stores all 4 bytes starting from R0
- To the memory address pointed to by R1

```
Memory[0x2000] = 0xDD
Memory[0x2001] = 0xCC
Memory[0x2002] = 0xBB
```

Memory[0x2003] = 0xAA

- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

```
B = 1: Transfer a byte (8 bits)
```

> Example

```
LDRB R0, [R1]
```

- Only loads 1 byte at 0x1000 → 0xAA
- The other 24 bits in **R0** are cleared

```
R1 = 0x1000

Memory[0x1000] = 0xAA

Memory[0x1001] = 0xBB

Memory[0x1002] = 0xCC

Memory[0x1003] = 0xDD
```

- LDR/STR Instruction Encoding
- > Byte/Word bit (B)

B = 1: Transfer a byte (8 bits)

> Example

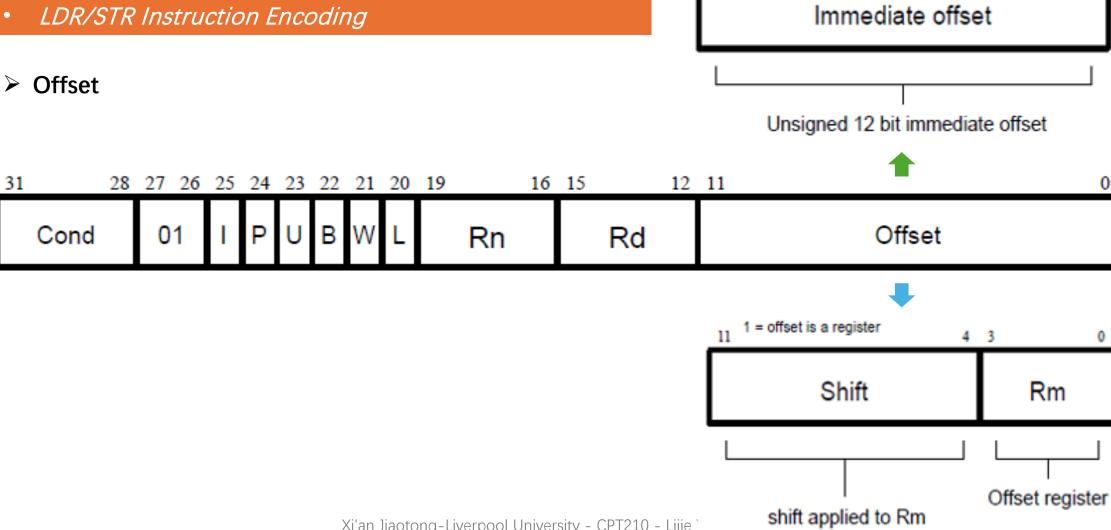
STRB R0, [R1]

- $R0 = 0 \times AABBCCDD$
- R1 = 0x2000

- Store only the lowest 8 bits (1 byte) of R0
- To the memory address pointed to by R1

Memory[0x2000] = 0xDD The rest of R0 is ignored





0 = offset is an immediate value

0



Shift Rm

Shift Offset register

shift applied to Rm

- LDR/STR Instruction Encoding
- > Offset
- LDR and STR support adding/subtracting offsets stored in another register:

```
LDR/STR{cond} Rt, [Rn, ±Rm{, shift}]

LDR/STR{cond} Rt, [Rn, ±Rm{, shift}]!

LDR/STR{cond} Rt, [Rn], ±Rm{, shift}
```

> Examples

Symbols + and – are not supported in VisUAL. They can be used in Keil IDE

```
LDR r0, [r1, r2, LSL #2]

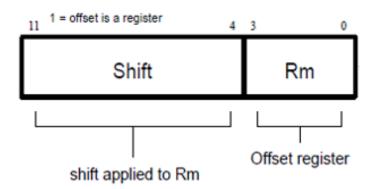
LDRB r0, [r1, r2, LSR #1]!

LDR r0, [r1], r2, LSL #3
```



- LDR/STR Instruction Encoding
- > Examples

- Type: Pre-indexed with register-based shifted offset
- Meaning:
  - Base address is in R1
  - Offset = R2 shifted left by 2 bits → R2 × 4
  - Effective address = R1 + (R2 ≪ 2)
  - Result is loaded into R0
  - R1 is not updated

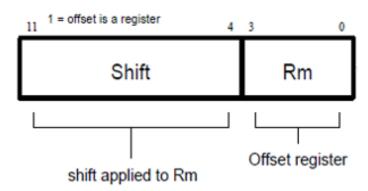




- LDR/STR Instruction Encoding
- > Examples

LDRB r0, [r1, r2, LSR #1]!

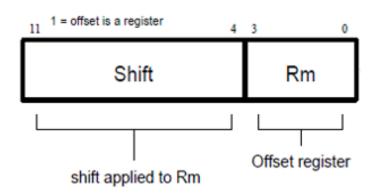
- Type: Pre-indexed with write-back, byte transfer
- Meaning:
  - Load 1 byte (because of LDRB)
  - Base address is in R1
  - Offset = R2 shifted right by 1 bit  $\rightarrow$  R2  $\div$  2
  - Effective address = R1 + (R2 >> 1)
  - Load byte from that address into R0
  - Then store the new address back into **R1** (because of !)





- LDR/STR Instruction Encoding
- > Examples

- Type: Post-indexed with register-based shifted offset
- Meaning:
  - Load 4 bytes (word) from address in R1 into R0
  - After loading, update R1 = R1 + (R2  $\ll$  3)  $\rightarrow$  R2  $\times$  8
  - The shift happens **after** the access (post-indexing)

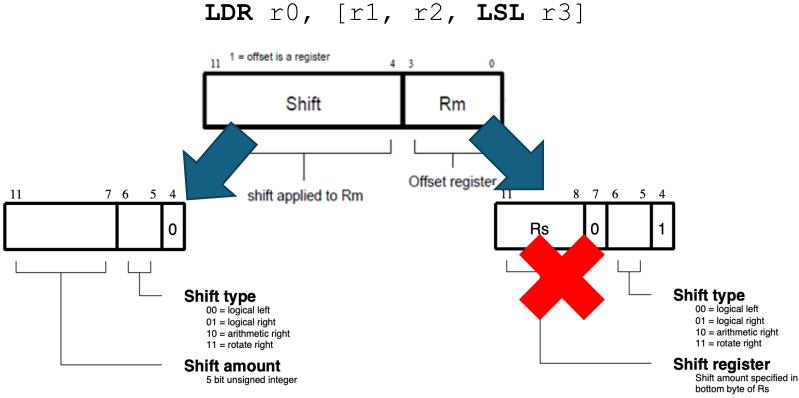




#### LDR/STR Instruction Encoding

Register-specified shift is **not** supported in LDR and STR.

Thus, you cannot write:





#### • LDR/STR Instruction Encoding

#### > Examples

What do the following machine code do?

	Cond	01	I	P	UВ	W	L	Rn	Rd	(	Offset	5
1	1110	01	0	1	10	0	0	0001	0000	0000	0000	0001
		-			-							
2	1110	01	0	1	10	1	1	0001	0000	0000	0000	0001
					•							•
3	1110	01	1	0	1 1	0	1	0001	0000	0011	0010	0010
		•	•	•	-	-						
4	1110	01	0	1	00	1	1	0001	0000	0000	0000	0001



Cond 01 I PUBWL Rn Rd Offset

# Answers

1110 01 01 10 0 0 0001 0000 0000 0000 0001

Step	Instruction
L is 0	STR
B is 0, transfer a whole word.	STR
Cond is 0b1110, which means "always".	STR
P is 1, pre-indexing	STR Rd, [Rn, ??]
Rn is 0b0001, which means R1.	STR Rd, [R1, ??]
Rd is 0b0000, which means R0.	STR R0, [R1, ??]
W is 0, do not update the base address (Rn)	STR R0, [R1, ??]
I is 0, the offset will be an immediate number.	STR R0, [R1, #?]
Offset is 1.	STR R0, [R1, #1]
$\ensuremath{\mathtt{U}}$ is 1, adding offset to base (offset is a positive immediate).	STR R0, [R1, #1]



Cond 01 I PUBWL Rn Rd Offset

# Answers

1110 01 01 1 0 1 1 0001 0000 0000 0000 0001

Step	Instruction
L is 1.	LDR
B is 0, transfer a whole word.	LDR
Cond is 0b1110, which means "always".	LDR
P is 1, pre-indexing	LDR Rd, [Rn, ??]
Rn is 0b0001, which means R1.	LDR Rd, [R1, ??]
Rd is 0b0000, which means R0.	LDR R0, [R1, ??]
W is 1, update the base address (Rn)	LDR R0, [R1, ??]!
I is 0, the offset will be an immediate number.	LDR R0, [R1, #??]!
Offset is 1.	LDR R0, [R1, #1]!
$\ensuremath{\mathtt{U}}$ is 1, adding offset to base (offset is a positive immediate).	LDR R0, [R1, #1]!



Cond 01 I P U B W L Rn Rd Offset

# Answers

1110 01 01 0 01 1 0001 0000 0000 0000 0001

Step	Instruction
L is 1.	LDR
B is 0, transfer a whole word.	LDR
Cond is 0b1110, which means "always".	LDR
P is 1, pre-indexing	LDR Rd, [Rn, ??]
Rn is 0b0001, which means R1.	LDR Rd, [R1, ??]
Rd is 0b0000, which means R0.	LDR R0, [R1, ??]
W is 1, update the base address (Rn)	LDR R0, [R1, ??]!
I is 0, the offset will be an immediate number.	LDR R0, [R1, #??]!
Offset is 1.	LDR R0, [R1, #1]!
U is 0, subtracting offset to base (offset is a negative immediate).	LDR R0, [R1, #-1]!



Cond 01 I P U B W L Rn Rd Offset

# Answers

1110 01 1 0 1 1 0 1 0001 0000 0011 0010 0010

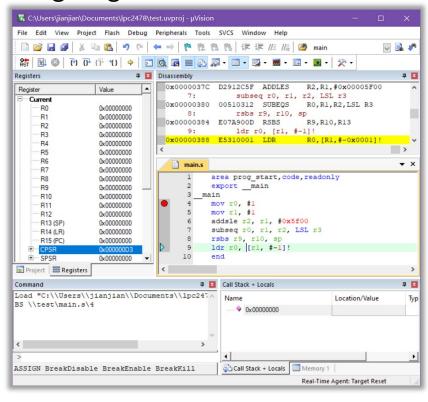
Step	Instruction
L is 1.	LDR
B is 1, transfer a byte.	LDRB
Cond is 0b1110, which means "always".	Shift type  00 = logical left 01 = logical right 10 = arithmetic right 11 = rotate right
P is 0, post-indexing	LDRB Rd, [Rn], ??
Rn is 0b0001, which means R1.	LDRB Rd, [R1], ??
Rd is 0b0000, which means R0.	LDRB RO, [R1], ??
₩ is 0, do not update the base address. (ignored as Post-indexing always updates the base address)	LDRB R0, [R1], ??
I is 1, the offset will be a register.	LDRB R0, [R1], Rm, ?????
Offset register is 0b0010 (R2), shift type is LSR (01), shift amount is 0b00110 (=6)	LDRB R0, [R1], R2, LSR #6
U is 1, adding offset to base (offset is a positive immediate).	LDRB R0, [R1], R2, LSR #6

If U is 0: LDRB R0, [R1], -R2, LSR #6 This syntax is only available on Keil.



# More Examples

- The encodings on the right are obtained in Keil IDE.
- VisUAL might give errors.



str r0, [r1, #1]	0xE5810001
str r0, [r1, #1]!	0xE5A10001
str r0, [r1], #1	0xE4810001
ldr r0, [r1, #1]	0xE5910001
ldr r0, [r1, #1]!	0xE5B10001
ldr r0, [r1], #1	0×E4910001

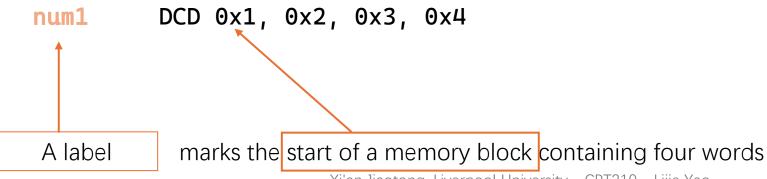
str r0, [r1, r2, LSL #1]	0xE7810082
strb r0, [r1, r2, LSR #2]!	0xE7E10122
str r0, [r1], r2, ASR #3	0xE68101C2
Idrb r0, [r1, r2, ROR #4]	0xE7D10262
ldr r0, [r1, r2, LSL #5]!	0xE7B10282
ldrb r0, [r1], r2, LSR #6	0xE6D10322

- ADR
- ➤ What is a label in ARM?

In assembly language, a label is a name you give to a memory location.

It can refer to:

- A block of data (e.g., constants, arrays)
- A line of code (e.g., function entry or a jump target)



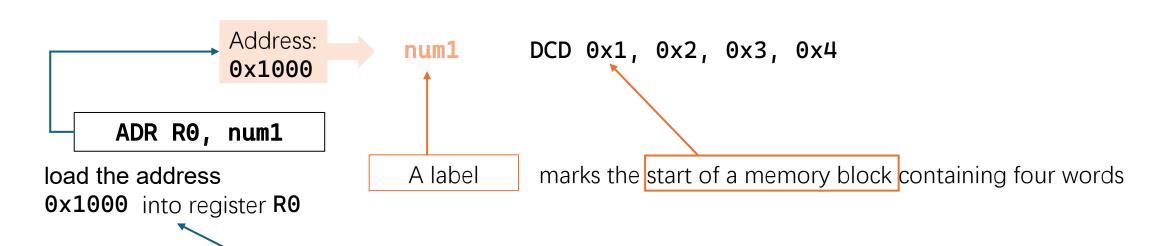


#### ADR

#### What is ADR?

ADR stands for **Address Register**.

It is an instruction to load the address of a label (memory location) into a register.



You're not loading the value 0x1, you're loading the address where 0x1 is stored

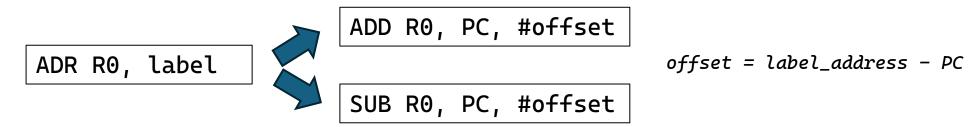


- ADR
- > Syntax

#### ADR Rd, label

- Rd: Destination register
- label: A named memory location (usually defined in .data or .rodata section)
- What does ADR actually do?
- It does not load the value at the label,
   but rather loads the address of the label into the register.

- ADR
- How does ADR work interbally?
  - ADR is implemented as either an ADD or SUB instruction under the hood.



- It uses PC-relative addressing (because it adds/subtracts an offset to/from the PC)
- The assembler calculates the distance between current PC and label and encodes it

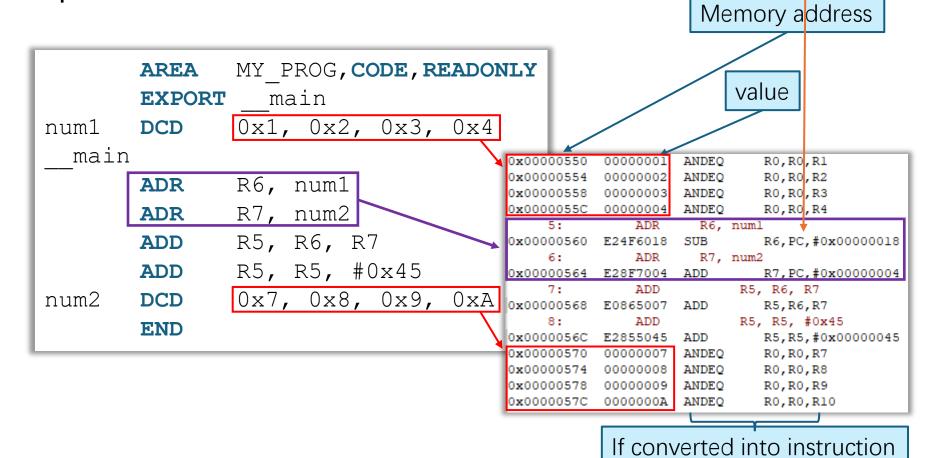
A basic rule in ARM:

PC = current instruction address + 0x00000008

Symbols will be converted into PC-relative values by assemblers

• ADR

> Example





- ADR
- > Try by yourselves

What do the following instructions mean?

Instruction	Meaning
ADR R0, label	?
LDR R0, =label	?

• LDR Pseudo-Instruction

LDR R0, =0xAABBCCDD

pseudo-instruction

- ARM instructions are 32-bit wide
- But you want to load a **full 32-bit constant**
- There isn't enough room in one instruction to include a full 32-bit number

LDR Pseudo-Instruction

LDR R0, =0xAABBCCDD

pseudo-instruction

#### What does the assembler actually do?

- 1. It places the constant (0xAABBCCDD) somewhere **near your instruction**, in a region called the **literal pool**
- 2. It then converts the pseudo-instruction into a real instruction like:

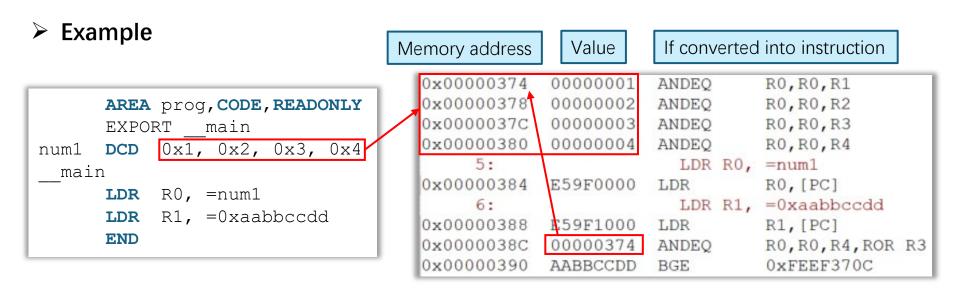
LDR R0, [PC, #offset]

This loads the value from **memory near the PC**, not directly from the instruction.



• LDR Pseudo-Instruction

LDR =something is a pseudo-instruction



- When the PC register is involved in instructions like LDR and MOV, its value is the address of the current instruction + 8.
- For instruction: LDR R0, [PC]
  - The address of the next instruction: 0x388
  - 0x388 + 0x4 = 0x38C, which points to the value 0x374 (address of num1)
- This design is to preserve compatibility for programs written for early ARM processors.



- LDR Pseudo-Instruction
- ➤ ? Why does LDR R0, =num1 give R0 = 0x00000374 (the address of num1) instead of loading the value stored at that address (e.g., 0x00000001)?

**✓** The short answer:

Because LDR R0, =num1 loads the address of the label num1, not the value stored at that label. It's a design behavior of the LDR pseudo-instruction with a label.



- LDR Pseudo-Instruction
- Why does LDR R0, = num1 give R0 =  $0 \times 000000374$  (the address of num1) instead of loading the value stored at that address (e.g.,  $0 \times 000000001$ )?
- What actually happens:

When you write LDR R0, =num1, the assembler creates an entry in the literal pool

0x0000038C 0x00000374

Address Value

00000001

00000002

00000003

00000004

E59F0000

E59F1000

00000374

AABBCCDD

ANDEO

ANDEO

ANDEO

ANDEQ

LDR

LDR

BGE

ANDEO

R0, R0, R1

RO, RO, R2

R0, R0, R3

RO, RO, R4

RO, [PC]

R1, [PC]

0xFEEF370C

R0, R0, R4, ROR R3

LDR R1, =0xaabbccdd

LDR RO, =num1

0x00000374

0x00000378

0x0000037C

0x00000380

5:

0x00000384

0x00000388

0x0000038C

0x00000390

6:

the CPU reads the value store at  $0 \times 0000038C$  (=  $0 \times 00000374$  = address of num1) and load it to R0



- LDR Pseudo-Instruction
- > Take-away information

- LDR =something is a pseudo-instruction
- The assembler creates a **literal pool** after the instruction section
- The PC is always current address + 8 when executing an instruction
- The offset is calculated so that the **LDR** accesses the right constant



# Before End: Final Exam & Assessment Explaination



# Final Exam

- Final Exam will take 70% of the module mark.
- Close-book exam: 0 materials will be provided, 0 materials can be brought in.
- On-site exam in Computer Lab Rooms from 10:00 am 12:00 pm (should arrive before 9:30 am) on May 26<sup>th</sup>
- 2 hours exam, conduct on LMO, with SEB (single-way exam, cannot go back to the previous questions, cannot leave the current page, otherwise the machine will be locked automatically).
- Questions are exacted from lectures and labs (about 9 big questions)
- Missing final exam = missing 70% of the module mark



### Assessment 1&2

- Assessment 1 & 2 will be combined into one and take 30% of the module mark
- On-site online assessment that happens during Tutorial → You MUST come to the tutorial in person
- You need to complete this assignment individually
- Photography/Al agencies are not allowed





- Questions are exacted from lectures and labs (about 6 questions)
- Scheduled during a specific time of your tutorial sessions (about 40 minutes)
- Bring your own LAPTOPS with FULL batteries (not tablet, not cell phone)
- Missing assessment = missing 30% of the module mark
- NO RESIT OF ASSESSMENT





