ENGN2219/COMP6719 Computer Systems & Organization

Convener: Shoaib Akram shoaib.akram@anu.edu.au



Australian National University

Shoaib Akram

Lecturer, School of Computing, (Jan 2020 –) Ph.D., 2019

Teaching: Computer Microarchitecure (semester 2) **Interests:** Hardware/software interaction and performance analysis

It is an interesting time to learn and research about computer systems. Traditional semiconductor laws are breaking down. But the society needs more compute power and storage capacity: Big Data, AI/ML, communication needs, among others.

My current focus is on enabling fast access to large amounts of information (data) using emerging memory and storage devices.

Logistics

Course webpage: <u>https://comp.anu.edu.au/courses/engn2219/</u> Lectures (on the website)

- Lecture slides
- Lecture videos (I will do my own recording)
- Weekly problem sets (for your practice only, Not Graded)
- Key Ideas and summary

Policies (will be up shortly)

 General conduct, assignment groups/submissions, support, management, grading

Resources

- Frequently asked questions
- Writing design documents
- Stuff needed to finish the assignments

Piazza

I will use Piazza for all communication

- If you ignore Piazza, you will miss key announcements
 - Drop-in sessions, make-up lectures, problems, exercises, corrections, lecture timing (ENGN2218 conflict)
- Ask questions on Piazza first (most likely you will receive a response quickly)
- Post solutions to weekly problems, ask your classmates if you are on the right track
- Ask private questions on Piazza to instructors
- Students are added/dropped automatically

Tutorials/Labs

Labs are a critical component of this course (one every week) Handout will be posted on the website "Labs" before each lab First six labs

- In each lab, you will finish a sub-component of design assignment 1
- If you finish the first six labs, you will finish ~50% of the first assignment
- We only mark the week 2 3 labs to make sure you are making progress and to give you feedback (due Monday 6 pm week 4)
- Week 3 lab accounts towards 5 points (out of 100) for assignment 1
 Week 7 lab (after the teaching break)
 - Test your design project

Week 8 – 11 Labs

 C programming (not really about C, but about learning key computer systems concepts, more on this later)

Assessments

Two assignments (60%)

- CPU design assignment (30%)
 - 5% due on 6 pm, Monday of week 4
 - Full assignment due on 12 pm, Monday of week 8
- Programming assignment (30%)
 - Due date: Monday 6 pm, week 13

Final Exam (40%)

- I will release problem sets and exercises throughout the course for preparation
- Mock-up exams will be available during/after the break

Assignment Submission

Extensions will be granted on a per-request basis

Via Email: Shoaib.Akram@anu.edu.au

Assignment submissions are handled via Gitlab

- You will learn about it in the labs
- Make a habit of using Git properly
- Push often, always pull the latest

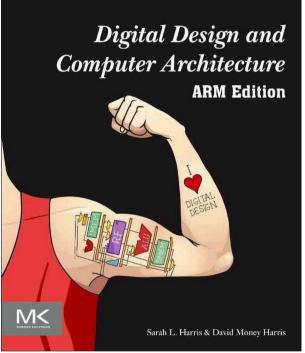
You can form group of up to two students to work on the assignment (one submission per group)

Ink/Whiteboard in Lectures

Try to take notes

- Help you think and most of the time I will do this is to solve a problem
- Ink + Wacom (not very stable)

Textbook

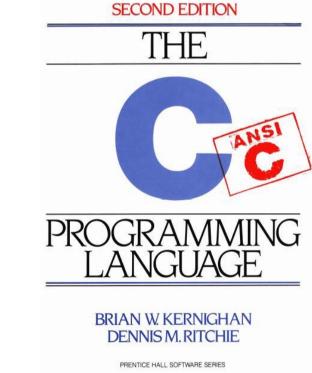


- Freely available online (check Piazza or course webpage)
- I will post the chapters/sections on the Lectures page after the lecture

Kernighan & Ritchie, The C Programming

Language, 2nd Edition

• "ANSI" (old-school) C





Council Bluffs, Iowa data center, Google (115, 000 sq. feet)



Self-flying nano drone 94 milli-watts

Research server for my students with special memory & storage devices



Fundamentals are important

All computer systems, big or small, have a few fundamental components

- Microprocessor (processor or central processing unit or CPU) for doing computation
- Main memory for storing temporary information and program data close to the processor
- Storage devices (e.g., disks or SSDs) for storing long-term or persistent information
- I/O devices to communicate with the external environment
 - Sensors
 - Peripherals

Most computer systems can be viewed as below

- Three key resources: CPU, memory, storage
- CPU is the heart of a computer system
- Processor can access memory much faster than storage



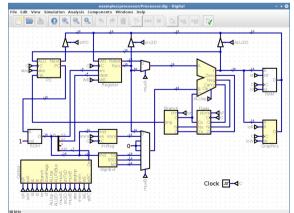
This Course is about ...

How does the general-purpose processor work? How do modern processors perform a wide variety of tasks?

How do processors interact with main memory and storage? How does the memory and storage system work?

The best way to learn how something works is to build one

- You will build a processor in Digital
- URL: https://github.com/hneemann/Digital



This Course is also about ...

A computer system is more than just hardware

- How does hardware and software interact?
- What should programmers know about hardware?
- C is a good vehicle for answering the above questions
- You can talk about hardware resources in high-level terms but still stay close to the hardware
- Key learning outcome of this course: How can you shoot yourself in the foot when writing C programs?
- Remember: It's not about C or Java or Python. It's about gaining a deep insight into computer systems!

A 5-Step Recipe for Failure

Stay out of the loop

- Do not check Piazza
- Do not ask questions
- And do not learn to use Gitlab Do not care what is going on in lectures

Do not attempt end of week problem sets

Okay, not every week

Do not come to labs OR do not read/attempt the lab handouts

And start the assignment on your own the weekend before due date (no way!)

Do not seek help from tutors

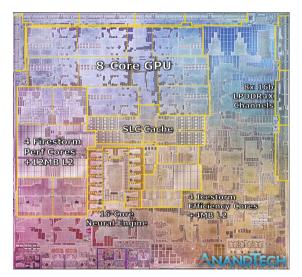
Do not communicate your problems to me!



How do engineers manage complexity?

- Look at components from a higher level
- Get into detail if necessary

No human (programmer) can track 10 billion elements. **Computer systems** work because of abstraction!

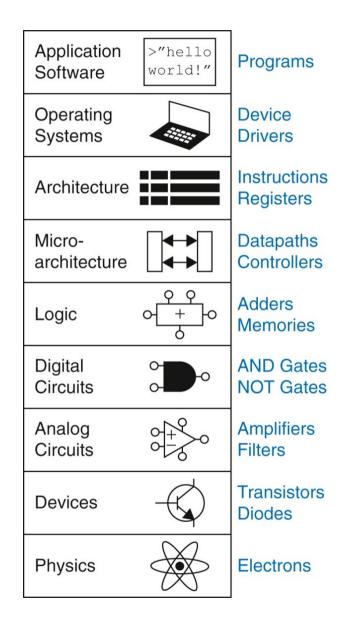


Apple M1 Chip Billions of transistors All working in parallel

Transformation Hierarchy

- We think of problems in English
 - Sort students by their UIDs
- The actual work is done by electrons
 - Do electrons speak English?
- How do we make the electrons do the work?
 - We use a systematic transformation hierarchy to transform the problem in English into electron movement
 - This transformation hierarchy is driven by our need to abstract away complexity

	Problem
	Algorithm
to	Program
	Architecture
	micro-arch
	circuits
	devices



Definitions

Abstraction: Hiding details to view the system from a high level

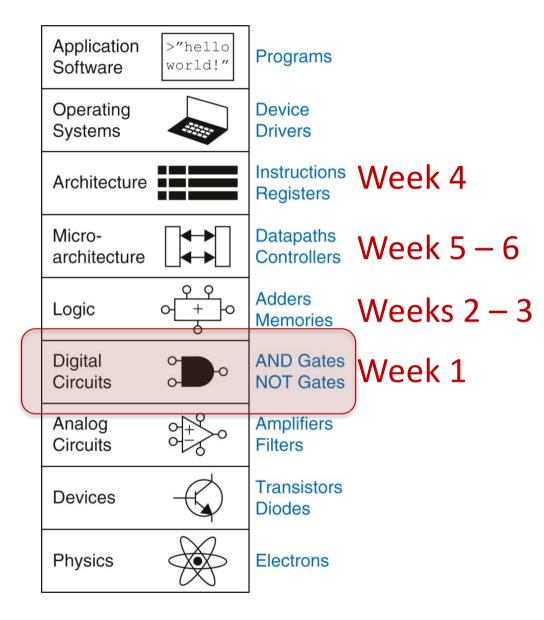
Deconstruction: Going from abstraction back to its component parts (breaking abstraction)

Low-level language: Languages that are tied to the machine architecture. Each architecture supports at least one low-level language called assembly.

High-level language: Programming languages that are at distance higher than the architecture. They are machine-independent. E.g., C, Java, Python, Rust, Ruby, Go

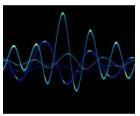
Instruction Set Architecture: A specification of all the instructions a processor can perform. Each instruction is an arithmetic (add, sub) or a data movement (fetch from memory) operation.

Microarchitecture: ISA has no physical significance. Microarchitecture is the physical implementation of an ISA.



Representing Information

Question: How many different values can each of these *physical* variables take?



Frequency of oscillation



Voltage on a wire



Temperature

Answer: Infinite

All these are continuous signals They contain infinite amount of information

Representing Information

Digital Systems: Represent information with discrete-valued variables, i.e., variables with a finite # distinct values

Modern digital systems use a **binary** (*two-valued*) representation



Binary Representation

Digital systems internally use "voltages" for representing binary variables

- → Low voltage means 0
- \rightarrow High voltage means 1



A **bit** is a unit of information. A binary variable represents one bit of information. To represent discrete sets with more than two elements, we combine multiple bits into a binary code

Binary Codes

Suppose we want to represent four colors: {red, blue, green, black}

- How many bits of information do I need?
- (00, 01, 10, 11)
- The assignment of the **2-bit binary code** to colors is *ad-hoc*
- Also legitimate is: (10, 11, 00, 01)

How many bits of information do I need to represent the alphabet set in English?

For 26 alphabets, we need 5 bits

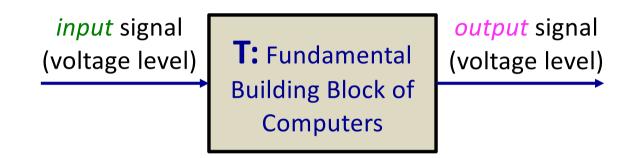
Information Content in a Binary Code

$D = Log_2 N bits$

The color set has four states: N = 4, # bits = 2 The alphabet set has 26 states: N = 26, # bits = 5

Conversely, If D is 2, N = 4 If D is 5, N = 32

Why do computers use binary?

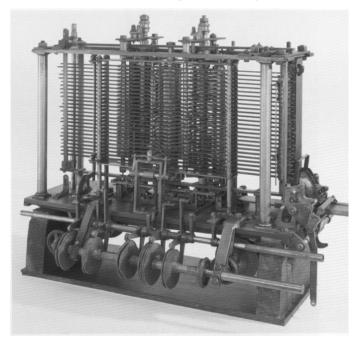


We can divide a continuous voltage range into ten levels to represent 0 – 9, but that would make **T** very complex

The fundamental building block of all computers is a transistor. A transistor can only distinguish two voltage levels. We call these voltage levels 1 and 0

Voltages and Transistors, Why?

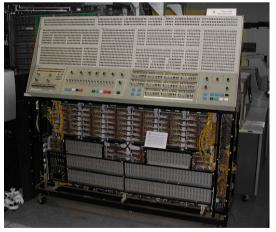
Mechanical parts: Not easy to scale to do large computations



The Analytical Engine Charles Babbage 1834 – 1871



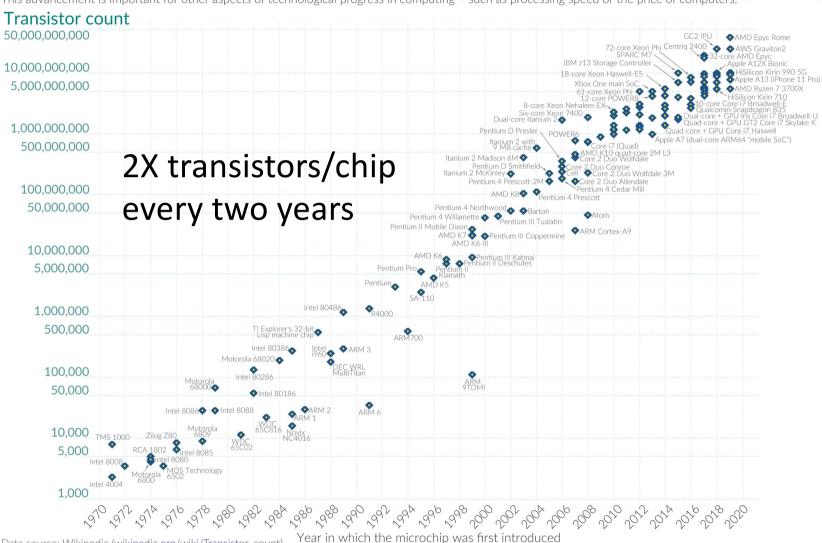
CDC 6600, 1964, \$ 2.5 M Slower than my phone



IBM 360, 1964



Apple M1, 2020 400 mm² 16 billion transistors



Moore's Law: The number of transistors on microchips doubles every two years

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Data source: Wikipedia (wikipedia.org/wiki/Transistor_count) Year in which the microchip was first introduced OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under

Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

TRUE and FALSE



F

FalseOffTrueOn

True and False are called logical values

- Logical variable is one that can be 1 or 0 (True or False)
- Boolean logic defines operations on logic variables

Our Plan

Presenting information to digital circuits

ightarrow Representing numbers as a string of 1's and 0's

ightarrow Number systems: to set a foundation for

efficient manipulation (add and subtract)



Operations on binary variables (1's and 0's)

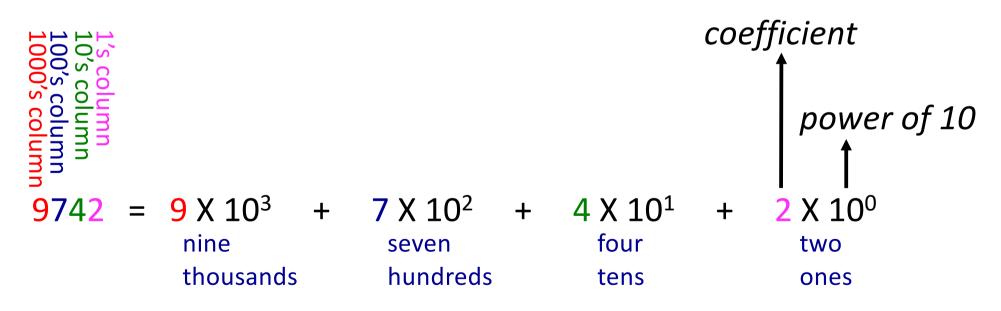
 \rightarrow Logic gates to perform operations on binary variables

Breaking the digital abstraction (self study)

 \rightarrow 1's and 0's as continuous physical quantities (voltage)

Decimal Number System

- Base 10 means 10 digits (0, 1, 2, 3, 4, 5, 6, 7, 8, 9)
- Multiple digits form longer decimal numbers
- Each column of a decimal number has 10 times the weight of the previous column



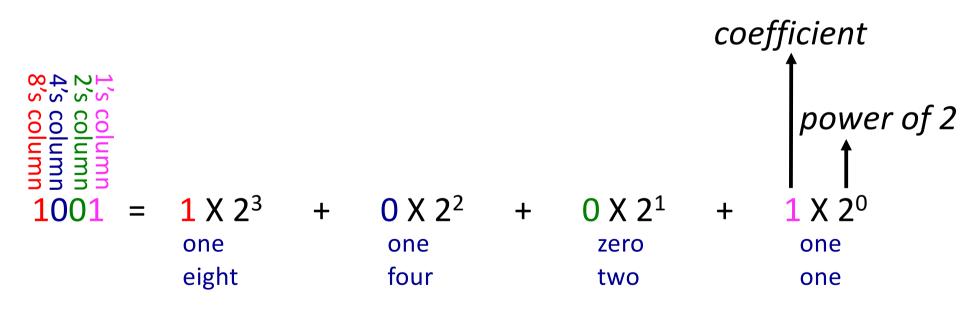
Range of Decimal Numbers

An N-digit decimal number represents one of $10^{\mbox{\tiny N}}$ possibilities

- 0, 1, 2, 3, ..., 10^N − 1
- 3 digits: 1000 possibilities in the range 0 999

Binary Numbers

- Base 2 means 2 digits (0, 1)
- Multiple bits form longer binary numbers
- Each column of a binary number has 2 times the weight of the previous column



Range of Binary Numbers

An N-bit binary number represents one of 2^N possibilities

- 0, 1, 2, 3, ..., 2^N−1
- 3 bits: 8 (= 2 X 2 X 2) possibilities in the range 0 7
- 4 bits: ?
- 5 bits: ?
- 10 bits: ?

Powers of 2

Columns #	Power of 2	Weight
0	2 ⁰	1
1	21	2
2	2 ²	4
3	2 ³	8
4	24	16
5	2 ⁵	32
6	2 ⁶	64
7	2 ⁷	128
8	2 ⁸	256
9	2 ⁹	512

Columns #	Power of 2	Weight	
10	2 ¹⁰	1024	Kilo
11	2 ¹¹	2048	
12	2 ¹²	4096	
13	2 ¹³	8192	
14	2 ¹⁴	16384	
15	2 ¹⁵	32768	
16	2 ¹⁶	65536	

Powers of 2

Power of 2	Decimal Value	Abbreviation
2 ¹⁰	1024	Kilo (K)
2 ²⁰	1048576	Mega (M)
2 ³⁰	1073741824	Giga (G)

What is 2²⁴ in decimal?

2²⁰ X 2⁴ = 1 M X 16 = 16 M

What is 2¹⁷ in decimal?

2¹⁰ X 2⁷ = 1 K X 128 = 128 K

Terminology

Byte

8 bits

Nibble

4 bits

Word

- Machine-dependent
- 8 16 bits (gadgets)
- 32 64 bits (high-end)

Most Significant Bit

1 0 0 0 0 0 0 1

The bit in the highest position

Least Significant Bit

1 0 0 0 0	0 0	1
-----------	-----	---

The bit in the lowest position

Terminology

Most Significant Byte

0 0 0 0 0 0 1 0 0 0 0 0 1

The byte in the highest position

Least Significant Byte
0 0 0 0 0 0 1 0 0 0 0 0 1

The byte in the lowest position

Decimal to Binary Conversion

Method # 1: Find the largest power of 2, subtract, and repeat

Example: Convert 53₁₀ to binary

53	32 X 1
53 – 32 = 21	16 X 1
21 – 16 = 5	4 X 1
5-4 = 1	1 X 1

2 ⁵	2 ⁴	2 ³	2 ²	2 ¹	2 ⁰
1	1	0	1	0	1

Decimal to Binary Conversion

Method # 2: Repeatedly divide by 2, remainder goes in the most significant position

Example: Convert 53₁₀ to binary

53/2	= 26	R: 1						
26/2	= 13	R: 0						
13/2	= 6	R: 1						
6/2	= 3	R: 0						
3/2	= 1	R: 1	2 ⁵	2 ⁴	2 ³	2 ²	2 ¹	2 ⁰
1/2	= 0	R: 1	1	1	0	1	0	1

Hexadecimal Numbers

Motivation: Tedious and error-prone to write long binary numbers

Hexadecimal or base 16: A group of four bits represent 2⁴ or 16 possibilities

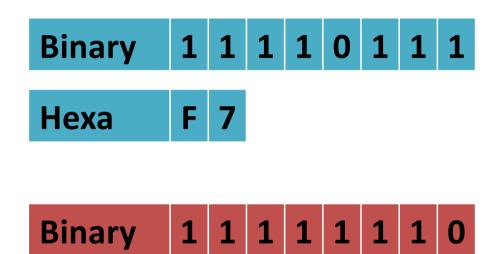
16 digits: 0 – 9, A, B, C, D, E, F

Column weights: 1, 16, 16² (or 256), 16³ (or 4096)

Hex Digit	Decimal Equivalent	Binary Equivalent
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
А	10	1010
В	11	1011
С	12	1100
D	13	1101
E	14	1110
F	15	1111

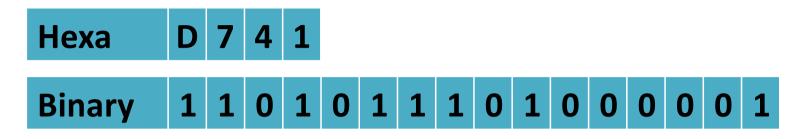
Hexadecimal Numbers

Binary to Hexadecimal

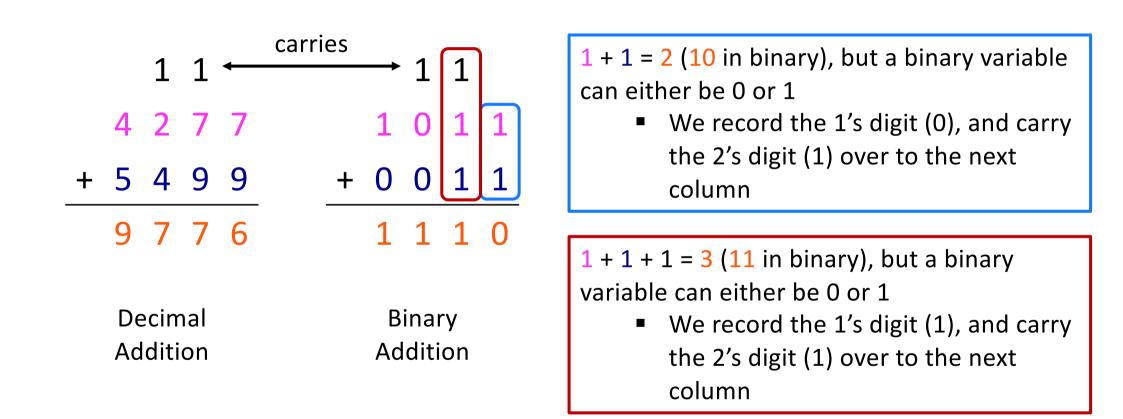




Hexadecimal to Binary



Binary Addition



Overflow

- Suppose we have two 4-bit numbers
 - If A = 1111 and B = 1111
 - What is A + B?
 - What is the largest value 4 bits can represent?
- Overflow
 - The result is too big to fit inside the available bits
 - We check the carry bit out of the most significant column to detect overflow

	1	1	1		
	1	1	1	1	15
+	1	1	1	1	15
1	1	1	1	0	30

Signed Binary Numbers

We need both positive and negative numbers to solve real-world problems

How do we make a string of 1's and 0's represent both positive and negative numbers?

If we write all possible combinations of O's and 1's in a disciplined fashion, maybe we can find a way

Most signi	ficant		least significant bit
bit		0	
	0	0	1
	0	1	0
	0	1	1
	1	0	0
	1	0	1
	1	1	0
	1	1	1

Signed Binary Numbers

Use the *most significant bit* to represent the sign: 0 means positive and 1 means negative

N bit sign/magnitude system: 1 bit for sign and N-1 bits for magnitude (absolute)

Signed Binary Numbers

What are the drawbacks of sign/magnitude Decimal representation? 0 0 0 +0+10 0 1 Ordinary binary addition does not work 0 1 0 +2 for sign/magnitude numbers What is the sum of +3 and -3 and +30 1 1 does the result make sense? 1 0 0 -0 1 0 1 -1 Awkward to have two different representations of the same number (0) -2 1 1 0

One's Complement

Tried in some early computers, Control Data
Corporation (CDC) 6600

Negative number: Take the representation of a positive integer and flip all the bits

Known commonly as 1's complement

 Same problems as the sign/magnitude representation Decimal

0	0	0	+0
0	0	1	+1
0	1	0	+2
0	1	1	+3
1	0	0	-3
1	0	1	-2
1	1	0	-1
1	1	1	-0

Two's Complement

A third system of representation for signed integers for which

- Ordinary addition works
- Single representation for zero

		Decimal					
	Α	?					
+	В	=	?	?	?	?	?
	С	=	0	0	0	0	0

		Decimal					
	Α	+5					
+	В	=	?	?	?	?	?
	С	=	0	0	0	0	0

		Decimal					
	Α	+5					
+	В	=	?	?	?	?	-5
	С	=	0	0	0	0	0

		Decimal					
	Α	+5					
+	В	=	1	0	1	1	-5
	С	=	0	0	0	0	0

What is the relationship between A and B?

		Decimal					
	Α	+5					
+	В	=	1	0	1	1	-5
	С	=	0	0	0	0	0

Some Observations

Observation # 1: If A + B = C, and A is +5, and C is 0, then B must be -5. (We have found a new representation for negative numbers.)

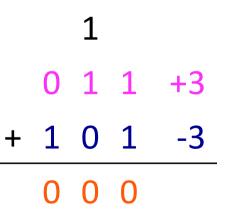
Observation # 2: To transform A to B (i.e., +5 to -5), we need to take 1's complement of A and then add +1. We say that B is **2's complement** of A

Observation # 3: Like sign/magnitude numbers, positive numbers have the MSB set to 0, and negative numbers have the MSB set to 1

Some Observations

Observation # 4: Ordinary addition works

- What is the sum of +3 and -3 in two's complement system, and *does the result make sense*?
- Since ordinary addition works, a circuit to add numbers can handle both addition and subtraction
 - Recall that, X A is equivalent to X + (–A)



2's Complement Circle 8888 <u>888</u> D S

More Observations

Observation # 5: There is only one representation for *zero*

Observation # 6: There is one more negative number than positive number

- With 3 bits, this number is 100
- With 4 bits, this number is 1000
- This negative number has no positive counterpart
- It is called the weird number
- The 2's complement of the weird number is the weird number (verify!)

2's Complement to Decimal

If MSB is 0

- It is a positive number. The magnitude is represented by the remaining N-1 bits
 If MSB is 1
 - It is a negative number. Take the two's complement of the (binary) number. The magnitude (of the negative number) is represented by the N – 1 bits

Practice and test your understanding using the two's complement circle

-	
5	
8-8	
 8 8	
8B	
8	
 8-9	
8-8 8-8	
8 8	
8	
N THE	
 8-28	
8-8	
8	
8-2-8	
8-8	
T.	
2	
8-1-1	
2	

Overflow in 2's Complement

- Suppose we have two 5-bit numbers
 - A = 01001 and B = 01011
 - What is A + B?
 - What is the largest value 5 bits can represent in 2's complement?
- Overflow
 - The result is too big to fit inside the available bits
 - Sum of two positive integers cannot produce a negative integer!

			1	1		
	0	1	0	0	1	+9
+	0	1	0	1	1	+9 +11
	1	0	1	0	0	-12

Overflow in 2's Complement

- Suppose we have two 5-bit numbers
 - A = 10100 and B = 11010
 - What is A + B?
 - What is the *smallest* value 5 bits can represent in 2's complement?
- Overflow
 - The result is too big to fit inside the available bits
 - Sum of two negative integers cannot produce a positive integer!

	1	0	1	0	0	-12
Ŧ	1	1	0	1	0	-6
	0	1	1	1	0	14

Overflow in 2's Complement

Observation # 1: If two number being added have the same sign bit and the result has the opposite sign bit (easy!)

Observation # 2: Unlike unsigned numbers, a carry out of the most significant bit does not indicate overflow

Observation # 3: The sum of a negative number and a positive number never produces an overflow (prove yourself!)

Range of Number Systems

Number System	Minimum	Maximum			
Unsigned	0	2 ^N -1			
Sign/Magnitude	-2 ^{N-1} + 1	2 ^{N-1} – 1			
Two's Complement	-2 ^{N-1}	2 ^{N-1} – 1			
<u>N = 3</u>	<u>N = 4</u>				
Unsigned: 0 to 7	Unsigned: 0 to ?				
Sign/Magnitude: -3 to 3	Sign/Magn	Sign/Magnitude: -? to ?			
2's Complement: -4 to 3	2's Complement: -? to ?				

Comparing Number Systems

Binary Representation	Decimal Value Represented								
		Signed	1's	2's					
	Unsigned	Magnitude	Complement	Complement					
000	0	0	0	0					
<mark>0</mark> 01	1	1	1	1					
<mark>0</mark> 10	2	2	2	2					
<mark>0</mark> 11	3	3	3	3					
100	4	-0	-3	-4					
101	5	-1	-2	-3					
110	6	-2	-1	-2					
111	7	-3	-0	-1					

Quiz: See	an	iy e	err	or	s? Ur	nsign	ed S	Signed	1' s	Con	np. 2'	s Con	np.
	0	0	0	0		0		0		0		0	
	0	0	0	1		1		1		1		1	
·	0	0	1	0		2		2		2		2	
	0	0	1	1		3		3		3		3	
	0	1	0	0		4		4		4		4	
	0	1	0	1		5		5		5		5	
	0	1	1	0		6		6		6		6	
	0	1	1	1		7		7		7		7	
	1	0	0	0		8		-0		-7		-1	
	1	0	0	1		9		-1		-6		-2	
	1	0	1	0		10		-2		-5		-3	
	1	0	1	1		11		-3		-4		-4	
	1	1	0	0		12		-4		-3		-5	
	1	1	0	1		13		-5		-2		-6	
	1	1	1	0		14		-6		-1		-7	
	1	1	1	1		15		-7		-0		-8	

Sign Extension

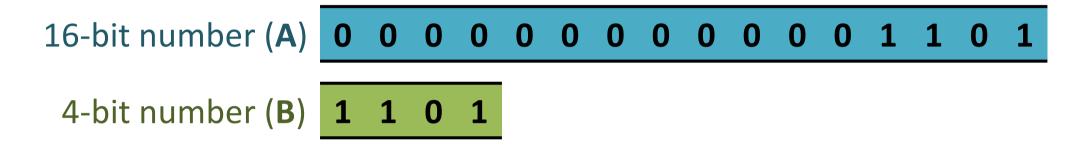
Question: What the difference between the 16-bit and 8-bit numbers below?

Answer: None. They both represent the positive number 5 *Leading zeros do not impact the magnitude of a binary number*

There are times when it is useful to allocate a small number of bits to represent a value

Sign Extension

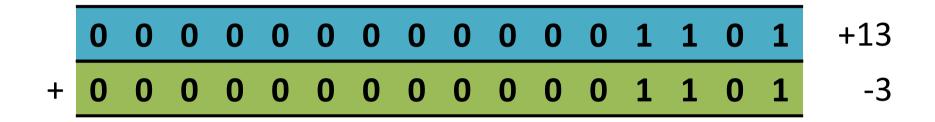
What value to the two numbers below represent?



What is the sum of **A** and **B**?

- Scenario # 1: Assume the absence of bits in B to be 0
- Scenario # 2: Assume the absence of bits in B to be 1

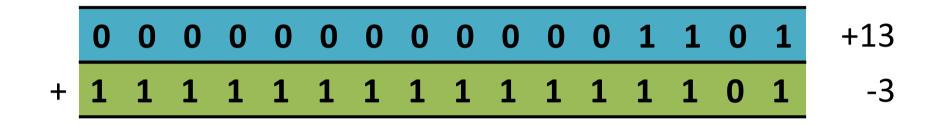




Х

The assumption that appending 0's will lead to correct addition was wrong

Scenario # 2



✓ The assumption that appending 1's will lead to correct addition was right

Sign Extension

Leading O's do not change the magnitude of the positive number

Leading 1's do not change the magnitude of the negative number

When a 2's complement number is extended to more bits, the sign bit must be copied into the most significant bit positions. We refer to the operation is Sign-EXTension or SEXT

Boolean Logic

A system of logic for binary variables

- Helps us reason about the interactions between two binary variables
 - X is 1 and Y is 0. What is $X \cap Y$?
- Formalizes key (simple) logical operations on binary variables
- Logical operations are the steppingstone for composing sophisticated operations (including arithmetic)

Logic Functions vs Gates

Logic gates are digital circuits that take one or more inputs and produce a binary output

- Logic gate is the physical realization of a logical function
 - The logic AND gate (*built with transistors*) implements the logical AND function
- The inputs are to the left, and the output is to the right
- The relationship between inputs and the output is described by a *truth table* or a *Boolean equation*

Truth Table

A convenient way to describe the behavior of logical functions

- Suppose A and B are input operands and Y is the output
 - A can be 0 or 1
 - B can be 0 or 1
 - A total of four combinations (rows)
 - Three columns (2 inputs and an output)

For the contrived example on the right, the Boolean equation for \mathbf{Y} is, $\mathbf{Y} = \mathbf{0}$

The values of A and B does not matter

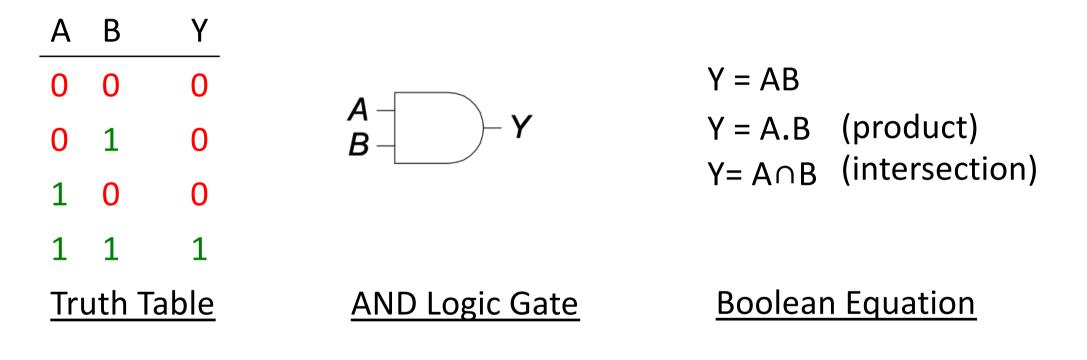
Α	В	Y
0	0	0
0	1	0
1	0	0
1	1	0

Truth Table with More Inputs

A	В	С	Y
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

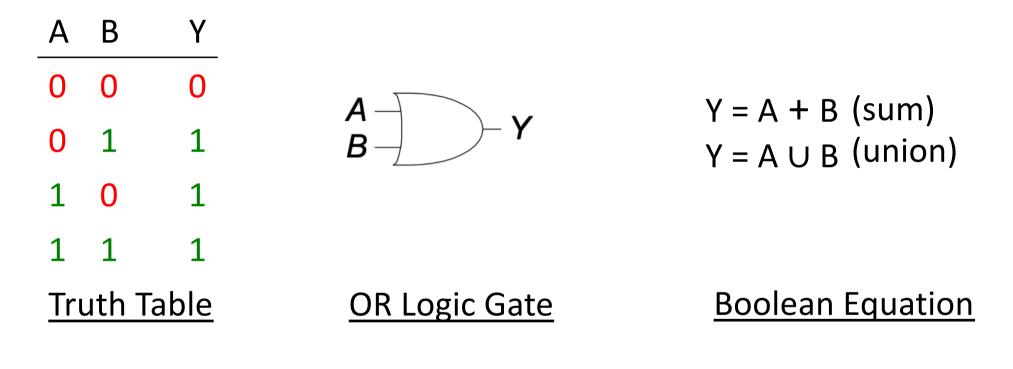
Boolean Equation for output Y: Y = 1**Note:** Soon we will see more interesting logic functions than Y = 0 and Y = 1

The AND Function

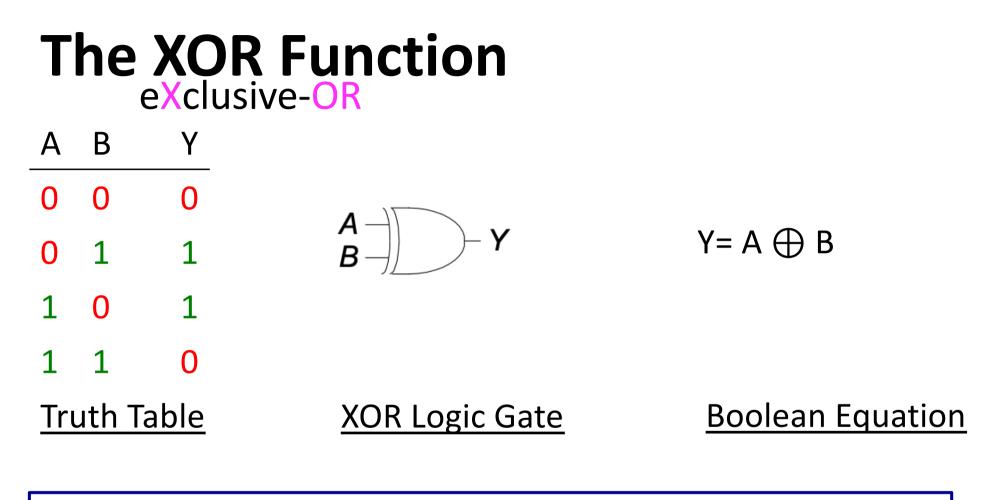


AND Function: The output Y is 1 if and only if both A and B are 1

The OR Function



AND Function: *The output Y is 1 if either A or B are 1*



AND Function: *The output Y is 1 if A or B, but not both, are 1*