Recall: what are Computers meant to do?

- We will be solving problems that are describable in English (or Greek or French or Hindi or Chinese or ...)
  and using a box filled with electrons and magnetism to accomplish the task.
  - This is accomplished using a system of well defined (sometimes) transformations that have been developed over the last 50+ years.
Recall:
Why use Binary and How to represent data in a computer?

- At the lowest level, a computer has electronic “plumbing”
  - Operates by controlling the flow of electrons
  - Electrons flowing on the wire when voltage exists

**Easy to recognize two conditions:**
1. presence of a voltage – call this state “1”
2. absence of a voltage – call this state “0”

**More complex to base state on value of voltage, but can be done**

**Think of the two states 0,1 as states of a switch**
- Change from 0 to 1 means throwing switch to turn on the light
- Presence of voltage on the wire means value of bit = 1 else 0

**Simple Switch Circuit**

- **Switch open:**
  - No current through circuit
  - Light is off
  - \( V_{out} \) is +2.9V

- **Switch closed:**
  - Short circuit across switch
  - Current flows
  - Light is on
  - \( V_{out} \) is 0V

**Switch-based circuits** can easily represent two states: on/off, open/closed, voltage/no voltage.

**A Quick review of some physics**
- Electricity corresponds to the flow of negatively charged particles called electrons. (see Ben Franklin)
- Particles of opposite sign, (+ve and –ve), attract each other
- Particles of the same sign repel each other.
- A voltage difference between 2 points captures the amount of work it would take to move charge from one point to another – analogous to an elevation difference in a waterfall
  - Current (like water itself), is the flow of electrons

**More Physics**
- Materials like metals are termed **conductors** because they allow the free flow of electrons
- Materials like rubber are termed **insulators** because they impede flow of electrons
- Resistors are devices that will conduct some current if you encourage the electrons with a potential difference
- Semiconductors are poor **conductors** and poor **insulators**, hence "semi." They can be used for either or both properties

**Ohm’s Law:** \( V = IR \)
**Voltage/Current and Electric Field**

- E-field produces "potential difference"  
  Aka: motivation for charge to flow  
  -  +  -  -  +  -  +  -  
  Direction of charge carrier (e-)

- Battery provides voltage  
  Aka: potential difference  
  +  -  +  -  +  -  +  -  
  Direction of current

- Ohm's Law: \( V = IR \)

---

**Switches to logic**

- A switch inherently represents two states, on/off  
  0-off  
  1-on  

- When put in a circuit, can start/stop current flow

- Putting multiple switches together, and we get basic logic structures

  **Switches are in series (AND)**
  - Both switches must be "on" for bulb to light up  
  (AND)

  **Switches are in Parallel (OR)**
  - Only 1 switch must be "on" for bulb to light up  
  (OR)

- Both switches must be "on" for bulb to light up
  (AND)

- Only 1 switch must be "on" for bulb to light up
  (OR)
Digital Circuits:
It's all about switching...

- Tubes
- Transistors
- CMOS FET

Computer use transistors as switches to manipulate bits
Before transistors: tubes, electro-mechanical relays (pre 1950s)
Mechanical adders (punch cards, gears) as far back as mid-1600s

Vacuum Tubes

- Also known as valves because they control the flow of electrons
  - Flow from Cathode to Anode
- First computer built using vacuum tubes

Historical Perspective

- ENIAC built in World War II the first general purpose computer
  - Used for computing artillery firing tables
  - 80 feet long by 8.5 feet high and several feet wide
  - Each of the twenty 10 digit registers was 2 feet long
  - Used 18,000 vacuum tubes
  - Performed 1,900 additions per second

Historical Fact: Do you know who are the “top secret rosies”?

Transistors

- Brought about a big change
  - Size
  - Speed
  - Precision
    - Moore’s law: they get smaller and faster
      - Can put more and more onto a single chip
- Also viewed in digital circuits as a “switch”
  - Transistors used in analog circuits
    - Stereos, recorders, Image proc., etc.
Transistor: Building Block of Computers

- Microprocessors contain millions of transistors
  - Intel Core Duo 2 (2006): 291 million
  - Intel 8-core Xeon Nehalem-EX (2010): 2.3 billion
  - nVIDIA GT200 (2008): 1.4 billion

- Logically, each transistor acts as a switch
- Combined to implement logic functions
  - AND, OR, NOT
- Combined to build higher-level structures
  - Adder, multiplexer, decoder, register, …
- Combined to build processor
  - LC-3 (LC4)

Basics of Digital Circuit Design

- How to build a switch?
  - Transistors
- How to build basic logic functions – gates using transistors?
  - Build simple gates (AND, NOT, OR, …) using transistors
- How to build more complex combinational logic using gates
  - Build Adders, multiplexer, decoder, storage devices using simple gates (AND, NOT, OR, …)
- Build a whole computer using complex logic devices
  - Assemble all the pieces together into an 'orchestra' – this is the CPU!

What is a transistor?

- A transistor is an electrical device that allows us to control the flow of current in a circuit
  - A transistor can act like an electronic “switch” in a circuit
  - A transistor can also function as an “amplifier” of voltage or current
- Over the decades, engineers have developed several electronic “switches” in circuits:
  - mechanical relays, vacuum tubes
  - diodes, transistors
  - MEMS devices, photonic, biological
- Switch-like behavior is important, because it can give rise to logic
  - In a CPU, we use transistors as switches, to implement logic gates

Transistor as electronic switch

- In the previous example with switches, someone must manually “flip” the switches to control the “input” to our gates
- In a computer we need a way to flip the switch by generating a signal
  - Transistor offers us this capability
  - We use voltage, to remotely flip the switch
- A transistor has 3 terminals:
  - This terminal is called “the drain”
  - Voltage applied to gate, allows current to flow between the drain and source
  - This terminal is called “the gate”
  - Terminal is called “the gate” controls the other two (using voltage)
How does a transistor work – Semiconductor basics
- Most materials are either insulators or conductors
  - They don’t “change” their properties
- Semiconductors: between insulator and conductor
- Semiconductors: Based on voltage applied to “gate” it is either insulator or a conductor
  - Electric field creates a circuit
  - Changes the device from an insulator to a conductor
- how does it work ??

Doping – not what you think
- We can improve the conduction of Silicon by doping it with other elements.
  - N-type regions are formed by adding small amounts of elements that have more than 4 electrons in their outer shell and, these extra electrons can serve as charge carriers.
  - N-type dopants – antimony (Sb), phosphorus (P), arsenic (As)

How does a transistor work?
- Begin at the beginning (what is it made of ?)
  - Currently transistors are etched on Silicon
    - Atomic symbol: Si – atomic number 14
  - In its crystalline state, silicon atoms form covalent bonds with four neighbors using their 4 outer electrons
  - At room temperature, Silicon is a semiconductor

P-type doping
- P-type materials are formed by adding elements that have 3 electrons in their outer valence shell.
  - These atoms create spaces in the lattice of covalent bonds into which electrons can flow.
- P-type dopants: boron (B), gallium (Ga), indium (In)
**Bottom Line**
- N-type materials are good semiconductors because they have extra electrons which are negatively charged and can be used to carry a current.
- P-type materials are good semiconductors because they have extra spaces into which electrons can move. These ‘holes’ can be thought of as positive charge carriers.

**A Diode (a pn-junction) – recall LED from lab**
- A union of P-type and N-type materials
- Functions as a one-way “valve” in an electric circuit
- Only allows current to flow in one direction

![Diode Diagram](image)

**Depletion region**
- Depletion region is an E-field that impedes the flow of current

**A Diode (a pn-junction)**
- Forward bias:
  - Depletion region gets smaller
  - Allows current to flow from + to -
  - Allows flow of electrons through junction
- Reverse bias (reverse the battery):
  - Depletion region gets bigger
  - Impedes flow of current from + to -
  - Impedes flow of electrons through junction

**Next up…the MOSFET (your 1st Transistor!)**
- MOSFET: Metal Oxide Semiconductor Field Effect Transistor
- Picture shows a cross section of such a device.
- Notice it has 4 electrical terminals: Source/Drain/Gate/Body
MOS FET (Metal Oxide SemiConductor)

- Source
- Gate
- Drain
- N-type
- P-type substrate
- Channel

MOSFET (your 1st Transistor!)

- MOSFET: Metal Oxide Semiconductor Field Effect Transistor
- Picture shows a cross section of such a device.
- Notice it has 4 electrical terminals: Source/Drain/Gate/Body

How we want it to work...

- Goal: Pass current through this device (from drain to source)
  - BUT we want to control that current (using the gate terminal)
  - If GATE is ON
    - electrons pass from source to drain through channel
  - If GATE is OFF
    - electrons cannot pass through channel
How we achieve this behavior...

- At "rest" we have (closed state)
  - 2 n-type spots (source/drain)
  - 1 p-type spot (channel region)
  - 2 back-to-back diodes!
    – Halts flow of electrons through channel (channel doesn’t exist!)

If we wish to turn device on:
- We apply a "positive" voltage to GATE with respect to BODY
  – This positive voltage "repels" holes from under the gate
  – "deletes" the future channel region of all its holes

If we go further:
- Apply a "very positive" voltage to the gate
  – Begins to attract electrons (from source & drain)
  – The channel region has been "inverted"
  – Connects (electrically) source and drain, so current can flow!
Two types of MOSFETs: nMOSFET and pMOSFET

• nMOSFET (nMOS): channel carries negative charges (electrons)
• pMOSFET (pMOS): channel carries positive charges (holes)

p-type MOS Transistor
• p-type
  • when Gate has positive voltage, open circuit between #1 and #2 (switch open)
  • when Gate has zero voltage, short circuit between #1 and #2 (switch closed)

Terminal #1 must be connected to +2.9V.

n-type MOS Transistor
• n-type complementary to p-type
  • when Gate has positive voltage, short circuit between #1 and #2 (switch closed)
  • when Gate has zero voltage, open circuit between #1 and #2 (switch open)

Terminal #2 must be connected to GND (0V).
Speed of MOSFET

- Dependent on many factors, 1 crucial factor: Length of Channel
  - Why? Electron takes less time to travel across smaller distance!
    - Currently, 11nm in length!

To turn on the light
What voltage do we apply here?

To turn the lightbulb off
To turn the lightbulb off:
Input A to switch = 0

A = 0
Switch - open

To turn the lightbulb ON:
Input to switch = 1

A = 1
Switch closed

Question: Where is the signal A=1 coming from?
.....generated by another circuit!!!

Light bulbs and computer hardware – what the &@??@&!

- Let’s look back at what we’ve learnt
  - Numbers can be represented as 0s and 1s
    - 1 is presence of voltage on line, 0 is no voltage on line
  - Arithmetic operations on these numbers
  - Logical operations on these numbers
- Starting point: how to implement the basic logic operators using transistors/switches?
  - NOT, AND, OR
- Next: how to implement arithmetic operations and other functions
  - Combinational circuits; example: adder
Logical Operations
- NOT, AND, OR, NAND, NOR, XOR
- These are binary functions
  - Input is binary, output is binary
- Boolean function – operates on boolean variables
  - Boolean function can be expressed using truth table
  - Eg: addition can be represented as a boolean function
- Recall from Discrete 1 - CS 1311: can implement any boolean function using AND, OR, NOT, etc.
  - In fact, can implement any bool function using just NAND
- Start by building these logical operator “gates” using transistors

Logic Gates
- Use switch behavior of MOS transistors to implement logical functions: AND, OR, NOT.
- Digital symbols:
  - recall that we assign a range of analog voltages to each digital (logic) symbol
  - assignment of voltage ranges depends on electrical properties of transistors being used
    - typical values for “1”: +5V, +3.3V, +2.9V

Ok….start building logic gates
- Use Complementary MOS (CMOS) circuits
- Using N type and P type transistors
- ‘signal’ is a 1 or 0 and nothing else
- Output value will be voltage measured at some point in the “circuit”
  - Need to determine where to designate the output point (i.e., where to measure)
- Inputs will be applied to the transistor gate
  - A line in the circuit always tied to 1 (voltage source) and one always tied to 0 (ground)
- Start by looking at the truth table for the logic function
- Important: The point (in circuit) where we measure voltage can only be a 0 or 1
  - During circuit design make sure this property holds
  - Else we have floating (unknown) values

So now what? How to go from “switch” to logic?
- Our first logic device will be an inverter: the NOT gate
- Logical Behavior: “inverts” the incoming signal:
  - Input: LOW-> output: HIGH
  - Input: HIGH->output: LOW

Truth Table
<table>
<thead>
<tr>
<th>INPUT</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW (0)</td>
<td>HIGH (1)</td>
</tr>
<tr>
<td>HIGH (1)</td>
<td>LOW (0)</td>
</tr>
</tbody>
</table>
How do we configure transistors to make an inverter?

We take advantage of the opposing nature:
- If pMOS turns on when GATE=0 Volts
- If nMOS turns on when GATE=High Voltage

Then if we put them together and connect their gates, we get inverting behavior!

IN=LOW (0 Volts)

IN=HIGH (2.9 Volts)

This configuration is called: CMOS

CMOS = Complimentary MOS Inverter

We have “jumped up” 1 level of abstraction
-- From transistors to “gate”
-- Technology inside the gate (CMOS here) isn’t as crucial as its behavior
-- Could be: transistors, vacuum tubes, biological device, etc.

3-D View of CMOS Inverter in Silicon

Note: we can make pMOS and nMOS transistors on the same piece of silicon

3-D of larger CMOS circuits

This is an SEM photo that shows all the metal interconnections on an IC.
pMOS/nMOS are at the very bottom.
Things to notice about a CMOS Circuit

- Uses both n-type and p-type MOS transistors
  - p-type
    - Attached to POWER (high voltage)
    - Pulls output voltage UP when input is zero
    - Call PMOS devices “pull up” devices
  - n-type
    - Attached to GROUND (low voltage)
    - Pulls output voltage DOWN when input is one
    - Call NMOS devices “pull down” devices

For all inputs, this configuration makes certain that output connected to GROUND or to POWER, but not both! (why?)

Circuit?

Truth table?

NAND Gate (AND-NOT)

Truth Table
AND Gate

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Add inverter to NAND.

Basic Logic Gates
- From Now On… Gates
  - Covered transistors mostly so that you know they exist
  - Note: “Logic Gate” not related to “Gate” of transistors
- Logic gates ~ Propositional logic operators
  - Propositional logic formula = Boolean logic circuit
- Will study implementation in terms of gates
  - Circuits that implement Boolean functions
- More complicated gates from transistors possible
  - XOR, Multiple-input AND-OR-Invert (AOI) gates

Truth Table for common 2 input gates

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>AND</th>
<th>OR</th>
<th>NAND</th>
<th>NOR</th>
<th>XOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Example: Your first combinational circuit
- Combinational logic circuits ~ propositional logic statements
- Use gates to implement the logic operators (‘functions’)
  - No necessity to show the circuit using transistors since each gate corresponds to an implementation using transistors
- Output = ((NOT A) AND B) OR C
- Need one AND gate and one OR gate (and one NOT gate/invertor)
More than 2 Inputs? Arbitrary Functions?
- AND/OR can take any number of inputs
  - \( \text{AND} = 1 \) if all inputs are 1
  - \( \text{OR} = 1 \) if any input is 1 (0 if all inputs are 0)
- Implementation
  - Multiple two-input gates or single CMOS circuit
- Can implement arbitrary boolean functions as a gate
  - More complex n- and p- networks

Visual Shorthand for Multi-bit Gates
- Use a cross-hatch mark to group wires
  - Example: calculate the AND of a pair of 4-bit numbers
    - \( A_3 \) is “high-order” or “most-significant” bit
    - If \( A \) is 1000, then \( A_3 = 1, A_2 = 0, A_1 = 0, A_0 = 0 \)

Shorthand for Inverting Signals
- Invert a signal by adding either
  - \( \overline{A} \) before/after a gate
  - A “bar” over letter

Some more observations about CMOS
- Note that when the circuit is fully ON or fully OFF there is no path from the high voltage to the low voltage so no current flows
- However, when the output is in the process of switching from one logic level to another, there can be overlap of the two switches being on
  - this causes a momentary short
    - (current goes from pwr-to-gnd)
  - Longer the short, more current you burn (more power wasted)!
- When current flows, device gets hot
  - The faster you switch the circuit, the more current flows, the more heat is generated, the hotter your laptop gets.
  - This has proven to be an important barrier to speeding up CMOS circuitry
  - led to Multi-Core processors….
The “Logic” Behind CMOS Gate Implementation

Transistors in series implement “AND”
- Current flows only if both are “ON”

Transistors in parallel implement “OR”
- Current flows if either is “ON”

CMOS is naturally inverting

Result: n-network implements function

NAND example
- n-network transistors in series gives AND
- Natural inversion gives NAND

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Gate Delays

- Which is the better implementation of 4-input AND?
  - One on the left
  - Why? It’s faster, 2 “gate delays” instead of 3

- Gate delays: longest path (in gates) through a circuit
  - Grossly over-simplified, ignores gate differences, wires
  - Good enough for our purposes

Gate Delay

- With any logic circuit there will be a short delay between the time you change one of the inputs and the time the output settles to its final value.
- This time is referred to as the gate delay.
- For modern circuitry, these gate delays are on the order of nano seconds ($10^{-9}$ seconds) or pico seconds ($10^{-12}$ seconds).
- Nonetheless, these delays ultimately limit the rate at which you can compute – limiting the number of operations you can perform per second.

The “Logic” Behind CMOS Gate Implementation

P-network is complement of n-network
- Series n-network → parallel p-network
- Parallel n-network → series p-network

NAND example
- p-network transistors in parallel
- Designing in CMOS:
  - We always design the n-network (aka – the pull-down network) first
  - Then, complement it and you’ve figured out the p-network (aka – the pull-up network)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Inverter Logic Gate
Reading
- Chapter 3 and Notes linked from webpage
- Start using Cedar Logic
  - If you have a Mac then use Logisim
- Review boolean algebra concepts from CS1311