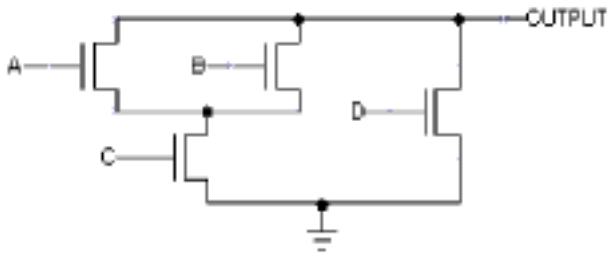
6.004 Computation Structures
Spring 2009

CMOS technology

Problem 1. The following diagram shows a schematic for the pulldown circuitry for a particular CMOS gate:



A. ★ What is the correct schematic for the pullup circuitry?

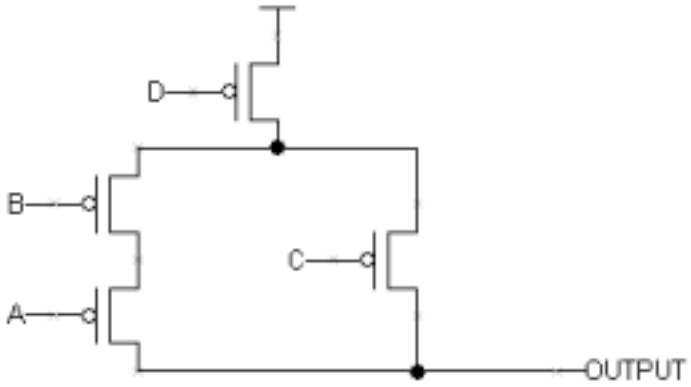
To figure out the complement of a FET circuit, decompose it into SERIES and PARALLEL subcircuits. In our example the NFET pulldown circuitry can be decomposed as:

PARALLEL(SERIES(PARALLEL(A,B),C),D)

The complementary circuit simply converts SERIES circuits into PARALLEL circuits and vice versa:

SERIES(PARALLEL(SERIES(A,B),C),D)

which we can then draw as a schematic using the complementary type of FET (in this example PFETs):



Note that series circuits can be drawn with the components in any order, so there are several

equally legitimate pullup schematics.

B. ★ Assuming the pullup circuitry is designed correctly, what is the logic function implemented this gate?

Start with the SERIES/PARALLEL decomposition of the pulldown given in the answer to the preceeding question:

PARALLEL(SERIES(PARALLEL(A,B),C),D)

Convert SERIES connections to AND and PARALLEL connections to OR:

OR(AND(OR(A,B),C),D)

Finally invert the answer to get

OUTPUT = -(D + C*(A + B))

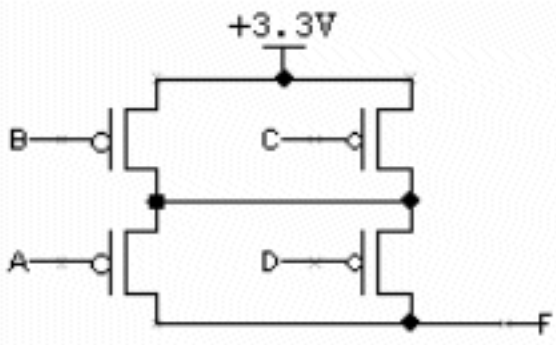
where "-" is NOT, "*" is AND and "+" is OR.

C. ★ Assuming the pullup circuitry is designed correctly, when the output of the CMOS gate above is a logic "0", in the steady state what would we expect the voltage of the output terminal to be? What would be the voltage if the output were a logic "1"?

When the output of a CMOS gate is a logic "0", we would expect the voltage to be essentially 0V. There is a miniscule amount of leakage through the pullup circuitry even though the PFETS are off, so the output isn't precisely 0V, but the current through the NFET pulldowns is many of order of magnitude larger.

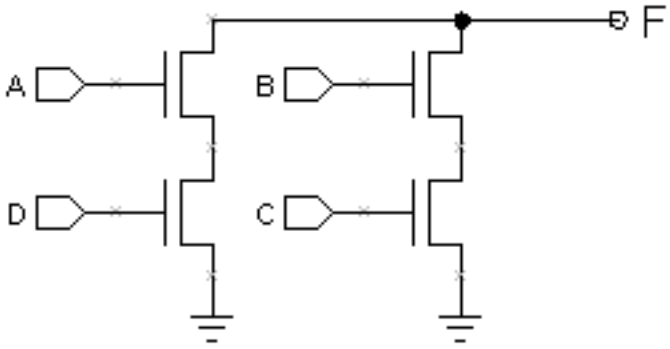
Similarly, when the output of a CMOS gate is a logic "1", we would expect the voltage to be the same as the power supply voltage (VDD).

Problem 2. The following diagram shows a schematic for the pullup circuitry for a particular CMOS gate:



A. ★ Draw a schematic for the pulldown circuitry for this CMOS gate.

The pullup is configured as SERIES(PARALLEL(A,D),PARALLEL(B,C)) so the pulldown should be the complement: PARALLEL(SERIES(A,D),SERIES(B,C))



B. ★ Assuming the pulldown circuitry is designed correctly, give an expression for the logic function implemented by this gate.

Using the pullup circuitry we can develop the following sum-of-products expression for F:

$$F = (\overline{A} + \overline{D})*(\overline{B} + \overline{C}) = \overline{A}*\overline{B} + \overline{A}*\overline{C} + \overline{B}*\overline{D} + \overline{C}*\overline{D}$$

Or we can look at the pulldown circuitry and express F as the complement of the function that describes when the pulldown is on:

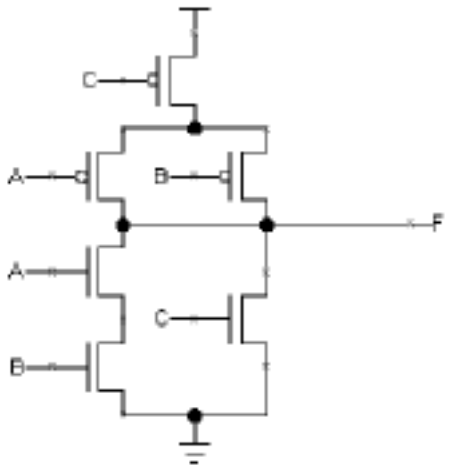
$$F = \overline{A*D + B*C}$$

Using several applications of DeMorgan's theorem this can be expanded to

$$F = (\overline{A*D})*(\overline{B*C}) = (\overline{A} + \overline{D})*(\overline{B} + \overline{C})$$

which is the expression we derived using the pullup circuitry.

Problem 3. Consider the following circuit built from nfets and pfets:



A. ★ Can this circuit be used as a CMOS gate? If not, explain why. If so, what function does it compute?

Since the pullup and pulldown circuits are complementary, the circuit is a legitimate CMOS gate. The function is

$$F = \overline{(A*B) + C} = \overline{(A*B)}*\overline{C} = (\overline{A} + \overline{B})*\overline{C} = \overline{A}*\overline{C} + \overline{B}*\overline{C}$$

B. ★ If we wanted the output voltage to change more quickly when going from a logic "0" to a logic "1", what changes would we make to the fets?

To increase the speed of a rising output transition, we need to charge the output node more quickly. The most effective change would be to increase the *width* of the PFET pullups which will increase the amount of pullup current they conduct when on.

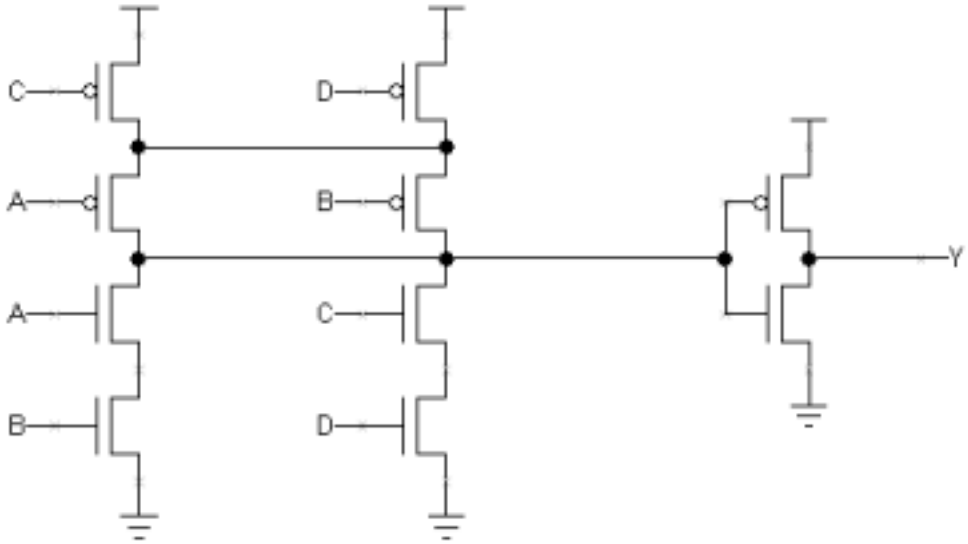
Note that increasing the width of the PFETS will also increase the capacitance of the output node since the PN junction capacitance of the PFET drain diffusions increases when the width of the PFET increases. But the speed gain from increased pullup current more than offsets the slowdown from increased capacitance.

We can also decrease the width of the two NFET pulldowns directly connected to the output

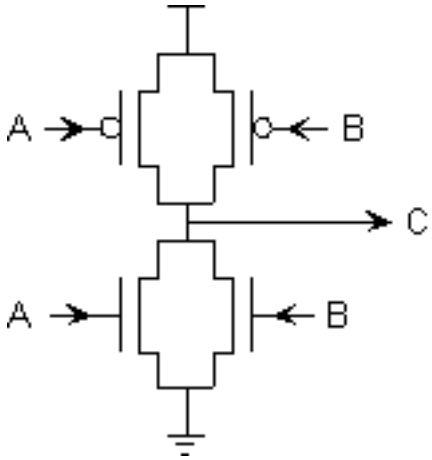
node. This will decrease the output capacitance and hence speed up any output transition.

Problem 4. Consider the 4-input Boolean function Y = (A*B) + (C*D) where "*" is AND and "+" is OR.

A. ★ Implement the function with a single 4-input CMOS gate and an inverter.



Problem 5. Anna Logue, a circuit designer who missed several early 6.004 lectures, is struggling to design her first CMOS logic gate. She has implemented the following circuit:



Anna has fabricated 100 test chips containing this circuit, and has a simple testing circuit which allows her to try out her proposed gate statically for various combinations of the A and B inputs. She has burned out 97 of her chips, and needs your help before destroying the remaining three. She is certain she is applying only valid input voltages, and expects to find a valid output at terminal C. Anna also keeps noticing a very faint smell of smoke.

A. What is burning out Anna's test chips? Give a specific scenario, including input values together with a description of the failure scenario. For what input combinations will this failure occur?

The chips are burning out when the pulldown and pullup are both active. This will occur when A=0, B=1 or when A=1, B=0.

B. Are there input combinations for which Anna can expect a valid output at C? Explain.

Yes, if A=1 and B=1, then C=0. Or if A=0 and B=0, then C=1

C. One of Anna's test chips has failed by burning out the pullup connected to A as well as the pulldown connected to B. Each of the burned out FETs appears as an open circuit, but the rest of the circuit remains functional. Can the resulting circuit be used as a combinational device whose two inputs are A and B? Explain its behavior for each combination of valid inputs.

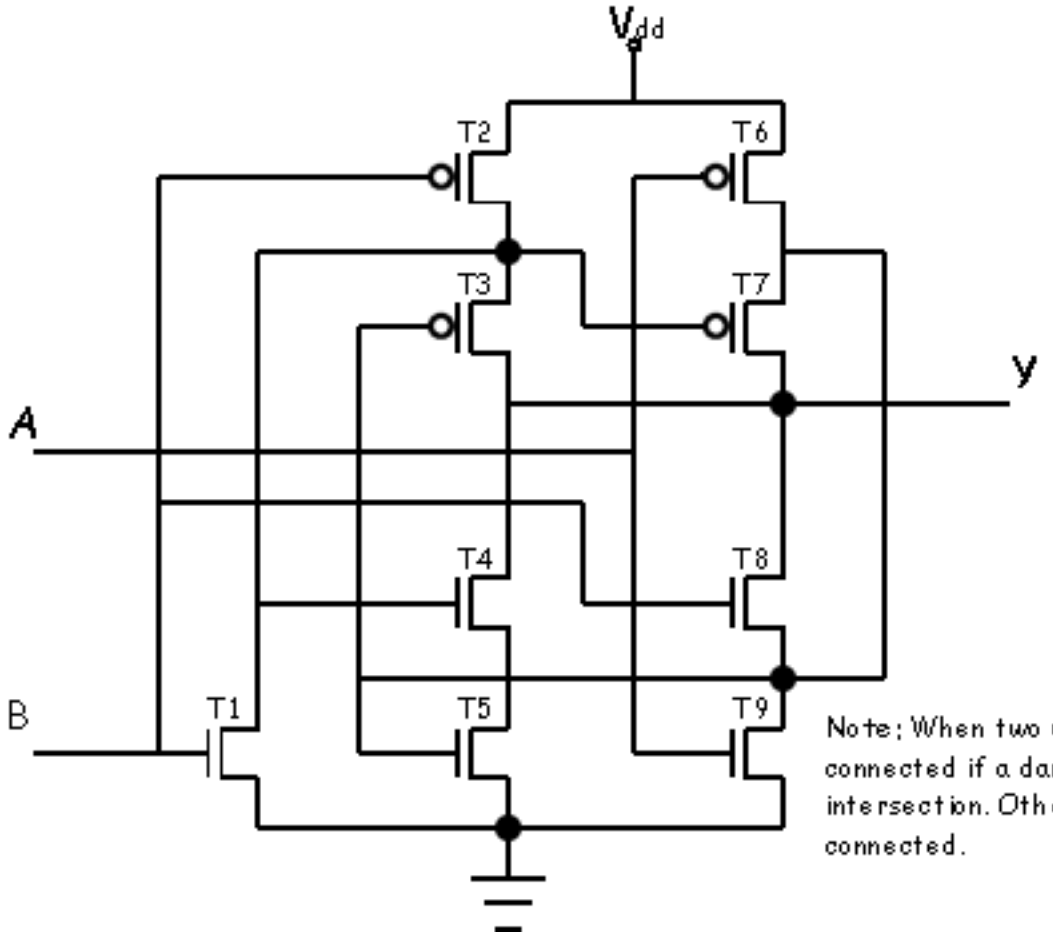
No. When A=1 and B=0, the circuit will burn out again, since the pullup and pulldown will be active, thus burning out the circuit. Also, the output is not defined when A=0 and B=1, since neither the pullup or pulldown are active.

D. In order to salvage her remaining three chips, Anna connects the A and B inputs of each and tries to use it as a single-input gate. Can the result be used as a single-input combinational device? Explain.

Yes. Since A=B, we are left with the following function (an inverter):

A	C
0	1
1	0

Problem 6. Occasionally you will come across a CMOS circuit where the complementary nature of the n-channel pull-downs and p-channel pull-ups are not obvious, as in the circuit shown below:



A. Construct a table that gives the on-off status of each transistor in the circuit above for all combinations of inputs A and B.

A	B	T1	T2	T3	T4	T5	T6	T7	T8	T9
0	0	off	on	off	**on**	**on**	on	off	off	off
0	1	on	off	off	off	on	**on**	**on**	on	off
1	0	off	**on**	**on**	on	off	off	off	off	on
1	1	on	off	on	off	off	off	on	**on**	**on**

B. Compute the output, Y, for each input combination and describe the function of the above circuit.

The output Y is connected to four pairs of transistors in series, so each of these pairs can affect the output.

when A=0 and B=0, transistors T4 and T5 are on, so Y=0

when A=0 and B=1, transistors T6 and T7 are on, so Y=1

when A=1 and B=0, transistors T2 and T3 are on, so Y=1

when A=1 and B=1, transistors T8 and T9 are on, so Y=0

Putting this together, we conclude that Y = XOR(A,B).

Problem 7. In lecture there was a brief overview of the CMOS fabrication process.

A. What keeps the source/drain diffusions of a MOSFET from shorting out to the substrate or to each other?

The source/drain diffusions are embedded in a substrate lightly doped to be of the opposite type. For example, NFETs have N-type source/drain diffusions embedded in a P-type substrate. A PN junction forms "automatically" where the source/drain diffusions and substrate come into contact. The designer of the circuit connects the substrate to the appropriate power supply rail (eg, GND or VDD) so as to guarantee that the PN junction is reverse-biased at all times. "Reverse-biased" means that V_P-V_N <= 0V. When the PN junction is reverse-biased, no current will flow across the boundary.

In the case of an NFET, the P-type substrate is tied to GND (V_P = 0), so even though the diffusion voltage V_N may anywhere between OV and VDD, V_P-V_N <= 0V.

B. Why does reducing the thickness of the thin oxide layer improve the performance of the mosfets?

An inversion layer (aka, the "channel") is formed in a FET when the gate voltage exceeds the threshold voltage of the device. The inverted channel conducts current between the source and drain diffusion. The depth of the channel and the strength of the inversion are proportional to the electric field generated by charges on the gate terminal -- the electric field is stronger if the gate terminal is brought closer to the channel by making the thin oxide thinner.

C. Why is silicon dioxide (SiO_2) deposited right before a new wiring layer is added to the surface of the wafer?

Silicon dioxide is an insulator, so by depositing it on the wafer we ensure that the new wiring layer is electrically isolated from the circuitry underneath.

D. How are connections between the wiring layers made?

Connections are made by etching holes in the insulating silicon dioxide layer before depositing and patterning a wiring layer. The wire material fills the holes making connections to the

underlying circuitry. In most modern processes, the holes are actually filled with tungsten, while the interconnect is formed from aluminum or copper.

E. If one wanted to *increase* I_{DS} for a NFET, how should it's dimensions be changed?

I_{DS} for a NFET is portional to W/L where W is its width and L its length. So to increase I_{DS} one should increase the NFET's width or decrease its length. Except in unusual circumstances, FETs are usually constructed using the minimum channel length allowed by the process, so decreasing the channel length is not an option.

F. Suppose there are two mosfets of width W and length L connected in parallel, i.e., all their terminal connections are identical. Given that I_{DS} of a mosfet is proportional to W/L, what would be the appropriate dimensions for a *single* mosfet that would have the same I_{DS} as the pair connected in parallel?

To first order, doubling the width of a mosfet will double its I_{DS}, so a single mosfet of of width 2W and length L would have the same I_{DS} as the pair connected in parallel.