2017 – Part IIA Module 3B2 – Integrated Digital Electronics
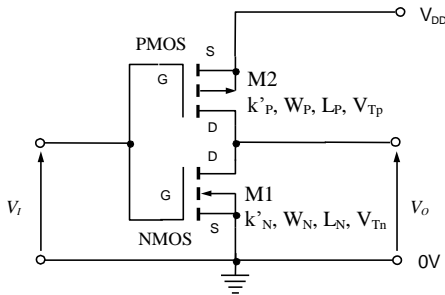
Qn 1 (a) (i) As defined V_{SP} is the point at which V_O and V_I are equal, and are expected to lie in the region somewhere between V_{DD} and 0 V.

At this point both MOSFETs are in the *saturation region*, and hence $V_{DS} > V_{GS} - V_T$

Hence the second of the given equations applies. By Kirchhoff, the currents I_{D1} and I_{D2} must be equal, hence:

$$\frac{k'_N}{2}\frac{W'_N}{L_N}\left(V_{SP}-V_{TN}\right)^2 = \frac{k'_P}{2}\frac{W'_P}{L_P}\left(V_{DD}-V_{SP}-V_{TP}\right)^2$$

Write $k_N = \dfrac{k'_N W_N}{L_N}$, and $k_P = \dfrac{k'_P W_P}{L_P}$; then

$$\frac{k_N}{2}\left(V_{SP}-V_{TN}\right)^2 = \frac{k_P}{2}\left(V_{DD}-V_{SP}+V_{TP}\right)^2$$

$$\sqrt{\frac{k_N}{k_P}}\left(V_{SP}-V_{TN}\right) = V_{DD}-V_{SP}+V_{TP}$$

Hence, $V_{SP} = \dfrac{\sqrt{\dfrac{k_N}{k_P}}\cdot V_{TN}+\left(V_{DD}-\left|V_{TP}\right|\right)}{1+\sqrt{\dfrac{k_N}{k_P}}}$

noting that V_{TP} is negative.

Let $X = \sqrt{\dfrac{k_N}{k_P}} = \sqrt{\dfrac{k'_N W_N L_P}{k'_P L_N W_P}}$, then $V_{SP} = \dfrac{X V_{TN}+\left|V_{TP}\right|+V_{DD}}{1+X}$

It can be seen that with a suitable choice of X (implying ratio of W_P/L_P to W_N/L_N), a range of different values may be obtained for V_{SP}, lying between V_{TN} and $V_{DD} - |V_{TP}|$. However, note that X is proportional to the square root of W/L, so V_{SP} is not a particularly sensitive function of k_N, k_P. [30%]

(ii) In the inverter, V_{SP} is normally chosen to be at around $V_{DD}/2$:

- to maximise the noise margins N_{MH} and N_{ML}

- this also matches the delay for rising and falling edges at the output, since the pull-up and pull-down conductances are matched.

The measures that have to be taken to achieve this can be seen by putting $V_{SP} = V_{DD}/2$ in the above equation:

$$\frac{V_{DD}}{2} = \frac{XV_{TN} + V_{TP} + V_{DD}}{1 + X} \quad \text{which we solve for } X$$

$$X = \frac{\frac{1}{2}V_{DD} + V_{TP}}{\frac{1}{2}V_{DD} - V_{TN}}$$

If we can assume, as is often the case, that: $\quad V_{TN} = |V_{TP}|$, then $\quad X = 1$.

Hence, $\qquad k'_N \dfrac{W_N}{L_N} = k'_P \dfrac{W_P}{L_P}$

i.e. the values of *W/L* must be chosen to be in the inverse ratio of the devices' k' (or of their mobility μ, assuming constant gate-oxide thickness t_{ox}, since $k' = \mu\varepsilon / t_{ox}$.

Note that the p-channel device is physically larger than the n-channel device, resulting in a loss of symmetry. [20%]

(b)   The RAM cell consists of a pair of cross-coupled inverters whose inputs may be accessed by means of switching transistors controlled by the **word** signal.
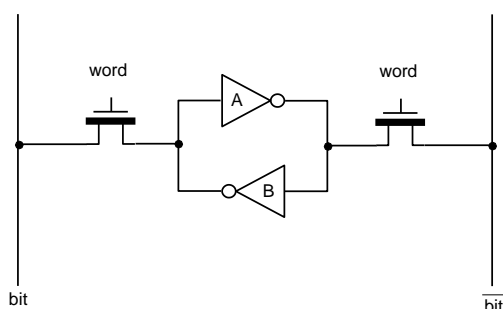
The cell can be in either of two stable states, corresponding to a stored logic '0' or logic '1'.

If inverter B is generating logic '1' that will cause A to output logic '0', which will cause B to output logic '1', so maintaining the original value.

The same argument hold is inverter B is generating logic '0', giving '1' at A's output, so maintaining this continuous set.

Hence the RAM cell may assume only two states and if isolated from external influences it will retain its state while power is applied.

This state may be changed, i.e. data *written*, by driving the vertical **bit** and **bit** lines to new complementary values, and operating the switching transistors by setting the word lines high. The bit-line signals override the signals currently being output by the inverters and switch them into a new state when necessary. It follows that the driver stages from which the input data is taken must have much greater drive capability, to obtain quick and decisive switching to the new state. Once this has been achieved, the word lines are reset to low, disengaging the switching transistors.

To read out data, both bit-lines are preset to precisely the same voltage, typically the mean of V_{low} and V_{high}. The lines are connected to sensitive comparators, which will initially indicate the equivalence of the signals on the lines. The switching transistors are then enabled. One inverter will drive its corresponding bit-line low, the other high – by a few millivolts. This is because the inverters are designed deliberately with low drive power ('weak'), and the bit-lines represent a substantial capacitive load whose charge is only slightly modified by the influence of the new incremental charges delivered by the two inverters.

The comparator output will then indicate logic '0' or logic '1', according to the sense of the perturbations introduced to the bit-lines by the RAM cell. [25%]

As implied, the RAM cell inverters must be sufficiently weak (by design) that the bit-line signals can override the signals currently being output by the inverters and switch them into a new state when necessary. If the driver stages from which the input is taken have much greater drive capability, quick and decisive switching will result.

To achieve the necessary 'weak' inverters, these are typically implemented using CMOS devices with low values of W/L, leading to low device conductances. [10%]

The additional circuit elements required are:

- Address decoder to select a specific horizontal row of cells to be manipulated
- Bank of sensitive comparators (alternatively called sense amplifiers). Design of these is a challenge since they must combine sub-mV sensitivity with high slew rate to read out the data fast.
- Column decoder, to select the output from a specific sense amplifier from the bank, and connect it to the output pin/s; also (for *write* operations) to connect the input data to the chosen set of bit lines.

Read/*write* control circuitry, which may allow the multiplexed use of a single pin for both input and output. [15%]

Assessor's Comments

The question was reasonably popular and while there were some extremely good results, a substantial proportion of answers were of rather disappointing quality. Descriptive parts in particular were poorly answered. Most candidates were able to set up an equation based on Kirchhoff's law to determine the switching point, but quite a few could not derive the simple solution that followed. Discussion of the measures needed to position the switching point was very variable. In the section on the static RAM, many candidates had difficulty explaining the salient features of its operation, though there were a few very good answers.

Qn 2.
(a) **Main reasons for popularity of CMOS**:

- Gate inputs are effectively open-circuit, very high impedance, easy to drive
- Power supply current for static/low frequency apps is very low, ideal for battery-powered portable devices
- Very good noise immunity, ~0.4 V_{DD} for inverter both low and high states (not quite so high for multi-input gates)
- Fully restored logic levels V_{DD} and 0V
- Can operate over wide range of supply voltages
- Creates little electrical noise
- Easily integrated with linear circuitry for complex mixed-signal designs

Main disadvantages:

- Not as fast as GaAs or some forms of bipolar
- Comparatively sensitive to static breakdown
- Liable to destructive latch-up as compound doped layers form thyristor-like structures [30%]

(b)
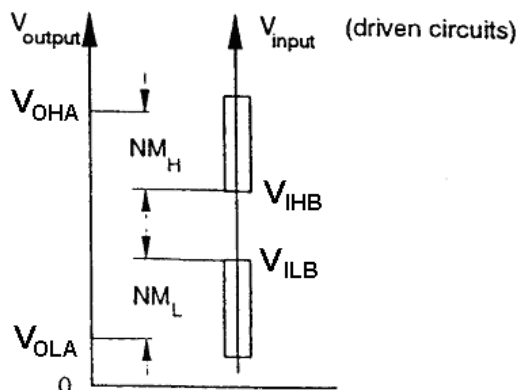
V_{OHA} and V_{OLA} represent respectively:
- the lowest voltage supplied by logic circuit A delivering logic '1', and
- the highest voltage supplied by an output delivering logic '0'

V_{IHB} and V_{ILB} represent respectively:
- the lowest input to B acceptable as logic '1'
- the highest input to B acceptable as logic '0'

If the output of A is connected to the input of B, the noise margins observed in the High and Low states are:

$$NM_H = V_{OHA} - V_{IHB} \qquad \text{and} \qquad NM_L = V_{ILB} - V_{OLA}$$

Both noise margins must be positive if the pair of circuits is to operate consistently. Their magnitude must be δ or greater if superimposed noise of voltage magnitude up to δ is to be rejected.

Note that the voltages V_{OH} and V_{OL} are liable to depend on the magnitude of the current flowing in the corresponding output devices – i.e. they depend on fan-out. [20%]

(c) (i) Bipolar drives CMOS

$N_{MH} = V_{OHbip} - V_{IHcmos} = 7.4 - 8.5 = -1.1$ V

$N_{ML} = V_{ILcmos} - V_{OLbip} = 1.0 - 0.4 = 0.6$ V

(ii) CMOS drives bipolar

$N_{MH} = V_{OHcmos} - V_{IHbip} = 9.9 - 2 = 7.9$ V

$N_{ML} = V_{ILbip} - V_{OLcmos} = 0.8 - 0.1 = 0.7$ V

(i)   would not work since $N_{MH} < 0$

(ii)   will work satisfactorily but note that it will be significantly more susceptible to noise in the low state.

[30%]

If 20 inputs were driven, the bipolar $\rightarrow$ CMOS margins (for operation assumed to be static or low-frequency) would be unaffected, since the table indicates that CMOS gates draw no input current.

The CMOS $\rightarrow$ bipolar noise margins are liable to be affected.

With the CMOS output high, the gate must source $210 \times 40 = 800$ μA when delivering '1' to the following 10 gates. This is not a particularly onerous demand.

With the CMOS output low, the gate must sink $20 \times 1.6 = 32$ mA at its output, from the following inputs. This might be beyond its capacity, or if not, it is liable to raise V_{OLcmos} and erode still further the already poor noise margin in the low state.

In either case, with 20 inputs driven, the additional capacitance will affect the rise/fall time achieved. Full analysis calls for more information about the devices in use and their dimensions, which determine the ability of the output stages to source/sink current to charge/discharge this parasitic load, but drivers implemented in CMOS are often limited cf. some bipolar families in terms of available output current.   [20%]

Assessor's Comments

This question was very popular, being attempted by almost all candidates and was in general answered to a good standard. Almost all were able to identify a satisfactory number of advantages of CMOS, and most could derive expressions for the noise margin, though a number used the wrong diagram. The problem part was done well, but a surprising number were unable to comment thoughtfully on the results obtained, or to state assumptions on which they were based.

3. (a) In a Mealy network, the primary outputs, Z=f(x,Q), is a function of primary inputs, x, and present states, Q, while in a Moore network, Z=f(Q) is a function of the present states only. A main difference compared to the Mealy network is that when a set of inputs is applied to the Moore network, the resulting outputs do not appear until after the clock pulse causes the flip-flops to change state. [10%]
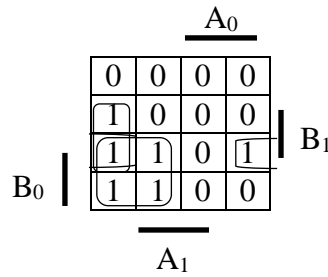
(b) It is a Moore configuration.   Z=f(Q)

	Z
x	Q
0	1
1	1

The output, $Z = x'Q + xQ = Q(x'+x) = Q$, depends on Q only.

[20%]

(c)   (i)

$A_0A_1B_0B_1$	x y z
0 0 0 0	0 1 0
0 0 0 1	1 0 0
0 0 1 0	1 0 0
0 0 1 1	1 0 0
0 1 0 0	0 0 1
0 1 0 1	0 1 0
0 1 1 0	1 0 0
0 1 1 1	1 0 0
1 0 0 0	0 0 1
1 0 0 1	0 0 1
1 0 1 0	0 1 0
1 0 1 1	1 0 0
1 1 0 0	0 0 1
1 1 0 1	0 0 1
1 1 1 0	0 0 1
1 1 1 1	0 1 0

K-map for x
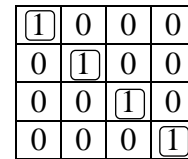
K-map for y
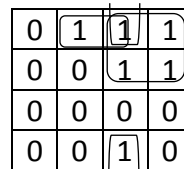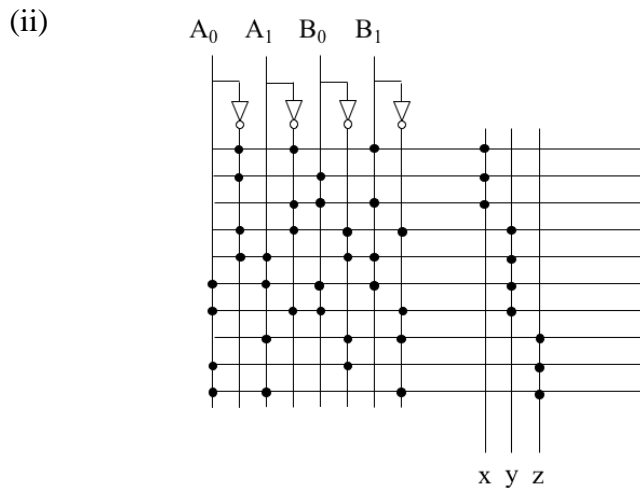

K-map for z

$x = \overline{A_0 A_1} B_1 + \overline{A_0} B_0 + \overline{A_1} B_0 B_1$

$y = \overline{A_0 A_1} \overline{B_0 B_1} + \overline{A_0} A_1 \overline{B_0} B_1 + A_0 A_1 B_0 B_1 + A_0 \overline{A_1} B_0 \overline{B_1}$
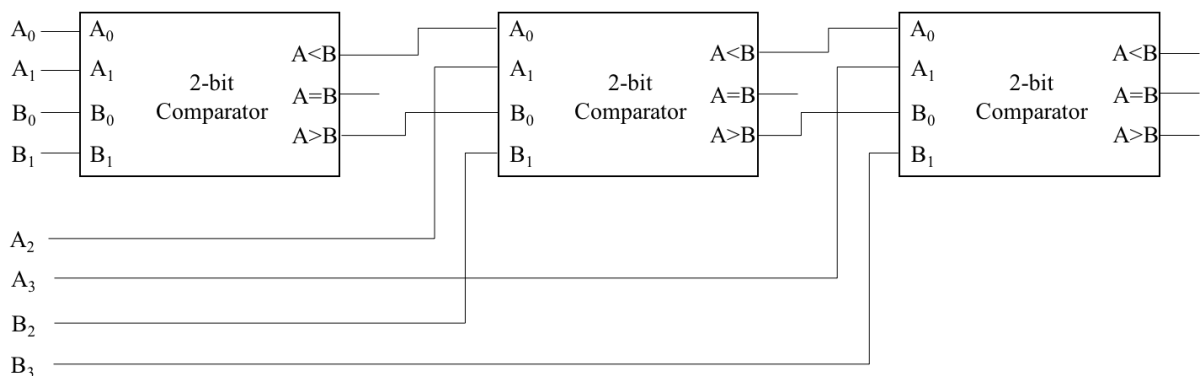
$z = A_1 \overline{B_0 B_1} + A_0 \overline{B_0} + A_0 A_1 \overline{B_1}$

[30%]

(ii)

The size of a PLA implementation (10x3 above) is comparable to that of a needed ROM (16x3). A ROM might offer a more economical solution. [20%]

(d)

We first compare the lower two bits of A and B, then next and next bits.

Example: Compare A=1011 and B=1010

We first compare 11 ($A_1 A_0$) and 10 ($B_1 B_0$) and we get a 1 at (A>B). Hence, on the next comparator, we put a 1 at B_0 and a 0 at A_0 and A_2 and B_2 at A_1 B_1 pins. So if A_2 is greater than B_2 then we get a 1 at (A>B) for 3 bits and if A_2 is less than B_2 we get a 1 at (A<B) for 3 bits and if A_2 is equal to B_2 then we compare A_0 and B_0. Similarly, we repeat to get the result. [20%]

Assessor's Comments

This question was very popular and well answered. Not all the candidates were able to identify the type of a sequential circuit, but almost all were able to implement combinational logic. Most of the candidates were able to understand how to choose between ROMs and PLAs.

4. (a) Multiplexers are suitable for single-output functions, preferably with few variables. ROMs are better for multiple-output functions but they become expensive and less efficient for very high number of variables. PLAs are good for multiple-output functions with lots of variables. They are expensive unless the number of variable is too high for a ROM. [20%]

(b) (i)
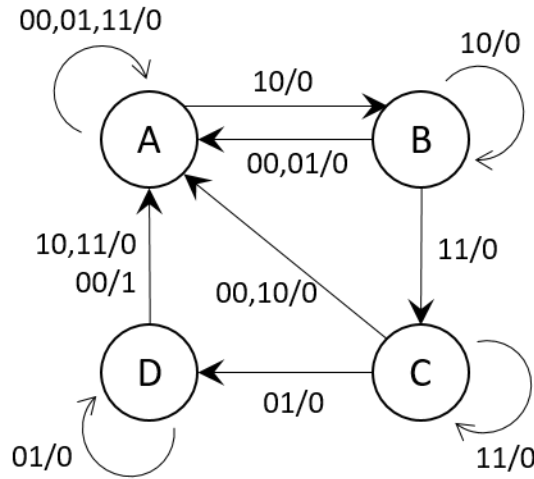
Present state		Next state for S_1S_2				Bistable inputs for $S_1S_2 = J_xK_x$				Bistable inputs for $S_1S_2 = J_yK_y$				Output			
	XY	00	01	11	10	00	01	11	10	00	01	11	10	00	01	11	10
A	00	00	00	00	01	0x	0x	0x	0x	0x	0x	0x	1x	0	0	0	0
B	01	00	00	11	01	0x	0x	1x	0x	x1	x1	x0	x0	0	0	0	0
C	11	00	10	11	00	x1	x0	x0	x1	x1	x1	x0	x1	0	0	0	0
D	10	00	10	00	00	x1	x0	x1	x1	0x	0x	0x	0x	1	0	0	0

[40%]

(ii)

$$J_x \quad K_x = S_1\overline{Y} + \overline{S_2}$$

$$J_y = \overline{X}S_1\overline{S_2}$$

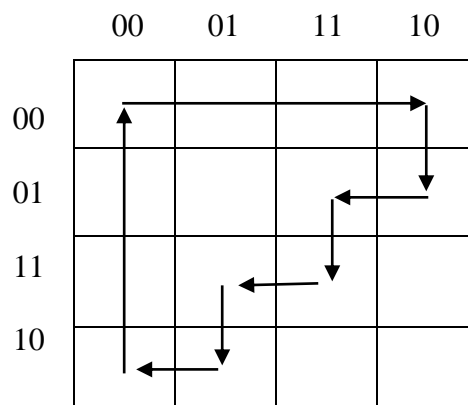$$K_y = X\overline{S_2} + \overline{S_1}$$

$$\text{Output} = X\overline{Y}S_1S_2$$
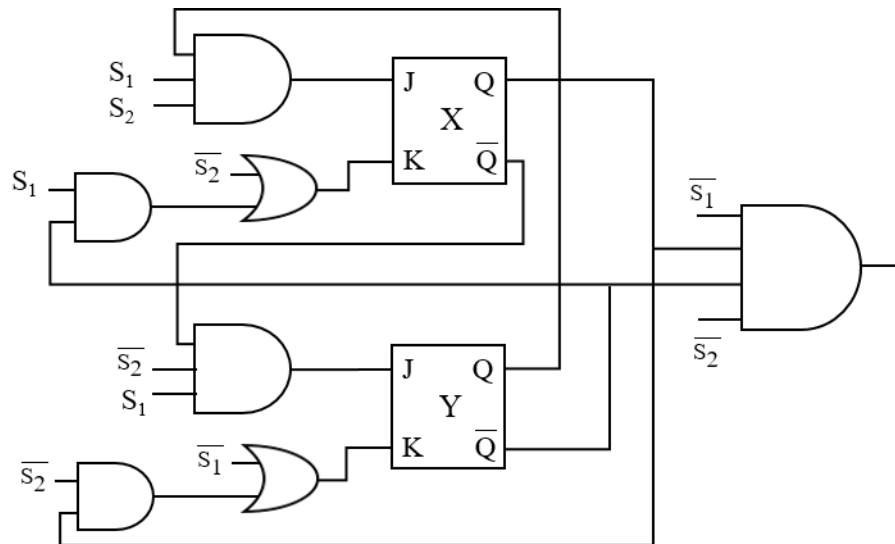
[20%]

(iii) The feedback from X and Y is delayed by few 10s ns. S_1 and S_2 cannot change that quickly since there is a considerable distance between the two detectors.

Only one bit changes at one time in the combined $S_1 S_2 XY$.

[20%]

Assessor's Comments

This question was attempted by 24 candidates and was answered to a good standard. The majority were able to identify the states of the sequential machine. Few were able to comment on the essential and functional hazards of the circuit.