Chapter 5

Solutions to Exercises

Exercise 5.1. Find two graphs that are structurally different but have the same degree distribution and clustering coefficients.

Solution. Here is one of many possible solutions:

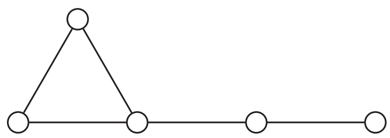
□

Exercise 5.2. Find the clustering coefficient of each vertex of the two graphs in Figure 5.2.

Solution. Neither graph has any triangles, so the clustering coefficient of each vertex in both graphs is zero. □

Exercise 5.3. Construct a graph that has two vertices of the same degree that have different clustering coefficients.

Solution. Here is one of many possible solutions:



The two vertices of degree 2 in the triangle have clustering coefficient 1, but the other vertex of degree two has clustering coefficient 0. □

Exercise 5.4. The clustering coefficients can be turned into a "distribution function," like what was done for the degree function, but with the minor difference that the domain must be the rational numbers, $\mathbb{Q}$. Namely, define $C\colon \mathbb{Q} \longrightarrow [0,1]$, where $C(p/q)$ is the fraction of nodes with clustering coefficient p/q. Write out this function for the graph in Figure 5.3.

Solution. This graph has seven vertices that have clustering coefficients 0, 4/15, 2/3, 2/3, 2/3, 1, and 1. The clustering coefficient distribution function is thus:

$$C(x) = \begin{cases} 1/7 & x = 0,\ 4/15 \\ 2/7 & x = 1 \\ 3/7 & x = 2/3 \\ 0 & \text{else} \end{cases}$$

□

Exercise 5.5. Calculate the betweenness of each node of the graph in Figure 5.3.

Algebraic and Discrete Mathematical Methods for Modern Biology. http://dx.doi.org/10.1016/B978-0-12-801213-0.00024-1

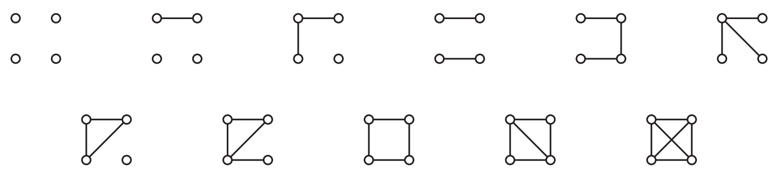
Solution. The node of degree 1 and the two nodes of degree 2 all have betweenness 0. For the "center fan node" of degree-3, exactly 1 of 16 shortest paths passes through it, so the clustering coefficient is 1/16. For the other two degree-3 nodes, exactly 1 of 17 shortest paths passes through it, so the clustering coefficient is 1/17.

To compute the betweenness of the center (degree-6) node, note that there are 5 shortest paths from the degree-1 node to one of the nodes on the "fan." Also, every pair of "fan nodes" a distance 2 apart have two shortest paths between them, one of which goes through the center. In all, there are 18 shortest paths between pairs of nodes, neither of which is the center, and 11 of them go through the center. The clustering coefficient of the center node is thus 11/18. □

Exercise 5.6. Calculate the betweenness of each node in the pentagon graph and in the heptagon (7-gon) graph.

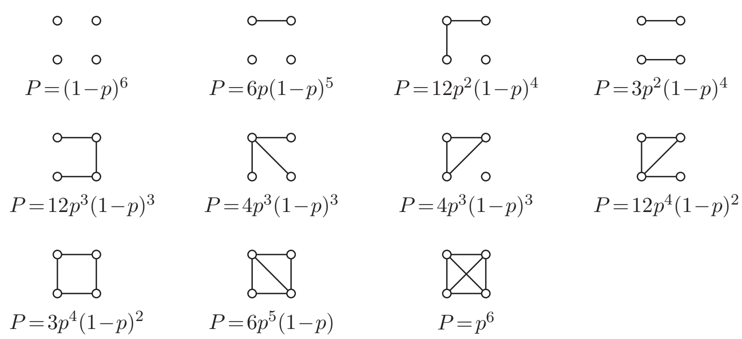
Solution. The clustering coefficient of each node in the pentagon graph is 1/6 and is 1/5 in the heptagon graph. □

Exercise 5.7. There are 11 distinct graphs on 4 vertices, which are shown here:



Compute the probability of each graph arising from the random graph model $G(4,p)$. Beware that most of these graphs can occur in multiple ways. For example, there are six different ways that the graph on one edge can arise.

Solution. The probabilities of all 11 graphs are shown below:



□

Exercise 5.8. There are 16 Boolean functions on two variables. Each one is determined by where it sends the four elements of $\mathbb{F}_2^2$: $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$, and so these 16 functions can all be arranged in a table as shown below. The first column represents the input $x=(x_1,x_2)$, and the other 16 columns represent the 16 possible Boolean functions. For example, the first two of these columns describe the two constant functions.

(x_1,x_2)																
$(0,0)$	0	1														
$(0,1)$	0	1														
$(1,0)$	0	1														
$(1,1)$	0	1														

Fill out the remaining 14 functions, and write each one in both its polynomial form and its Boolean form. Given a Boolean function $f(x)$, the two-column table consisting of the input $x=(x_1,x_2)$ and the output $f(x)$ is called the *truth table* of the function. Thus, the table represents the truth tables of all 16 Boolean functions on two variables.

Solution. Here are all 16 functions. Each one is labeled f_k, where k is the decimal equivalent of the four-digit binary number read down the column. The double vertical lines further group the functions by how many 1s are in their truth table. After the two constant functions, the next first four functions can all be constructed using a single AND statement:

$$f_1(x_1,x_2) = \text{and}(x_1,x_2) = x_1 \wedge x_2 = x_1x_2 \qquad f_2(x_1,x_2) = x_1 \wedge \overline{x_2} = x_1(1+x_2)$$
$$f_8(x_1,x_2) = \text{nor}(x_1,x_2) = \overline{x_1} \wedge \overline{x_2} = (1+x_1)(1+x_2) \qquad f_4(x_1,x_2) = \overline{x_1} \wedge x_2 = (1+x_1)x_2$$

(x_1,x_2)	f_0	f_{15}	f_1	f_2	f_4	f_8	f_3	f_5	f_6	f_9	f_{10}	f_{12}	f_7	f_{11}	f_{13}	f_{14}
(0, 0)	0	1	0	0	0	1	0	0	0	1	1	1	0	1	1	1
(0, 1)	0	1	0	0	1	0	0	1	1	0	0	1	1	0	1	1
(1, 0)	0	1	0	1	0	0	1	0	1	0	1	0	1	1	0	1
(1, 1)	0	1	1	0	0	0	1	1	0	1	0	0	1	1	1	0

The last four functions can all be constructed using a single OR statement:

$$f_7(x_1,x_2) = \text{or}(x_1,x_2) = x_1 \vee x_2 = x_1 + x_2 + x_1x_2 \qquad f_{11}(x_1,x_2) = x_1 \vee \overline{x_2} = 1 + x_2 + x_1x_2$$
$$f_{14}(x_1,x_2) = \text{nand}\,(x_1,x_2) = \overline{x_1} \vee \overline{x_2} = 1 + x_1x_2 \qquad f_{13}(x_1,x_2) = \overline{x_1} \vee x_2 = 1 + x_1 + x_1x_2$$

The last six functions consist of four functions one just one variable, and two functions that require both AND and OR:

$$f_3(x_1,x_2) = x_1, \quad f_{12}(x_1,x_2) = \overline{x_1} = 1 + x_1, \quad f_5(x_1,x_2) = x_2, \quad f_{10}(x_1,x_2) = \overline{x_2} = 1 + x_2$$
$$f_6(x_1,x_2) = \text{xor}(x_1,x_2) = (x_1 \wedge \overline{x_2}) \vee (\overline{x_2} \wedge x_1) = x_1 + x_2$$
$$f_9(x_1,x_2) = \text{xnor}(x_1,x_2) = (x_1 \wedge x_2) \vee (\overline{x_1} \wedge \overline{x_2}) = 1 + x_1 + x_2$$

□

Exercise 5.9. A Boolean function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is *symmetric* if $f(x_1,\dots,x_n) = f(x_{\pi_1},\dots,x_{\pi_n})$ for all permutations π of $\{1,\dots,n\}$. In other words, the output is only determined by how many entries of x are equal to 1.

(a) Determine which of the 16 Boolean functions on 2 variables (see previous exercise) are symmetric.

(b) How many symmetric Boolean functions are there on n variables?

(c) Find all symmetric Boolean functions on three variables. For each one, either identify it a commonly known functions (e.g., constant, AND, OR) or give it a simple name that accurately describes it.

Solution.

(a) Half of the 16 Boolean functions on two variables are symmetric: both constant functions, AND, OR, NAND, NOR, XOR, and XNOR.

(b) Every symmetric function f is uniquely determined by the following $n+1$ values: $f(0\cdots 0)$, $f(10\cdots 0)$, $f(110\cdots 0)$, …, $f(11\cdots 1)$. Because there are two choices for each, there are 2^{n+1} symmetric Boolean functions on n variables.

(c) There are $2^4 = 16$ symmetric Boolean functions on three variables, which are shown on the following table:

x	0	1	and	g_2	g_4	nor	maj	par	g_6	g_9	1 + par	min	or	g_{11}	g_{13}	nand
(0, 0, 0)	0	1	0	0	0	1	0	0	0	1	1	1	0	1	1	1
(1, 0, 0)	0	1	0	0	1	0	0	1	1	0	0	1	1	0	1	1
(1, 1, 0)	0	1	0	1	0	0	1	0	1	0	1	0	1	1	0	1
(1, 1, 1)	0	1	1	0	0	0	1	1	0	1	0	0	1	1	1	0

In addition to the familiar names, maj is the "majority" function that returns the most common bit, and min $= 1 +$ maj is the related "minority" function. The parity function par is just the sum modulo 2. Some of the other functions can be written in terms of these. For example, $g_6 = \text{nand} + \text{nor}$, and $g_9 = 1 + \text{nand} + \text{nor}$. □

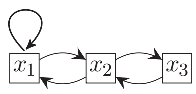
Exercise 5.10. Consider the Boolean network on three nodes whose update functions are

$$f_1(x_1, x_2, x_3) = x_1 \vee x_2, \quad f_2(x_1, x_2, x_3) = x_1 \vee x_3, \quad f_3(x_1, x_2, x_3) = x_2.$$

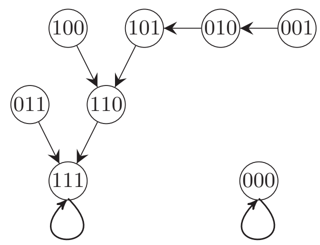
Write the functions in polynomial form. Then draw the wiring diagram and the phase space of this Boolean network. Identify the limit points and the transient points. Using the ADAM software (http://adam.plantsimlab.org/) will save considerable time and accuracy.

Solution. The polynomial form, wiring diagram, and phase space are shown here.

$$\begin{cases} f_1(x_1, x_2, x_3) = x_1 + x_2 + x_1x_2 \\ f_2(x_1, x_2, x_3) = x_1 + x_3 + x_1x_3 \\ f_3(x_1, x_2, x_3) = x_2 \end{cases}$$



Wiring diagram



Phase space

The phase space has two limit points: $(0, 0, 0)$ and $(1, 1, 1)$. The other six points are transient. □

Exercise 5.11. Let $(X, \mathcal{F})$ be a Boolean network with global update function $f: \mathbb{F}_2^n \to \mathbb{F}_2^n$. The set of transient states of f can be expressed formally as the set

$$\text{Trans}(f) := \{x \in \mathbb{F}_2^n \mid f^{(k)}(x) \neq x \text{ for all } k \geq 1\},$$

where $f^{(k)}(x)$ denotes f applied to x exactly k times. Complete the following statements with formal mathematical definitions.

(a) The set Per(f) of periodic points is:

$$\text{Per}(f) := \left\{x \in \mathbb{F}_2^n \mid \quad \right\}.$$

(b) The fixed points of f are the set:

$$\text{Fix}(f) : \left\{x \in \mathbb{F}_2^n \mid \quad \right\}.$$

(c) The garden-of-Eden state of f is the set:

$$\text{GoE}(f) := \left\{x \in \mathbb{F}_2^n \mid \quad \right\}.$$

Solution.

(a) The set Per(f) of periodic points is:

$$\text{Per}(f) := \left\{x \in \mathbb{F}_2^n \mid f^{(k)}(x) = x \text{ for some } k > 1\right\}.$$

(b) The fixed points of f are the set:

$$\text{Fix}(f) : \left\{x \in \mathbb{F}_2^n \mid f(x) = x\right\}.$$

(c) The garden-of-Eden state of f is the set:

$$\text{GoE}(f) := \left\{x \in \mathbb{F}_2^n \mid \nexists y \in \mathbb{F}_2^n \text{ such that } f(y) = x\right\}.$$

□

Exercise 5.12 (Exploratory)**.** Can you create a Boolean network on three nodes whose phase space consists of a single length-8 cycle? Can you create one whose phase space consists of a single chain leading into a fixed point?

Solution. Explore this by going to http://adam.plantsimlab.org/. □

Exercise 5.13. For each of the 16 Boolean functions from Exercise 5.8, determine if it is canalizing.

Solution. All the functions are canalizing except the two constant functions and $\text{xor}(x_1, x_2) = x_1 + x_2$, and its logical negation $\text{xnor}(x_1, x_2) = 1 + x_1 + x_2$. □

Exercise 5.14. The function $f(x, y, z) = x \vee y \vee z$ is nested canalizing, and all six orderings of the variables are nested canalizing sequences. Find a 3-variable NCF with fewer than six nested canalizing sequences. Can you find an NCF that has only one?

Solution. The function $f(x, y, z) = x \wedge (y \vee z)$ has only two canalizing sequences: $x < y < z$ and $x < z < y$. Every NCF with more than one variable will have at least two canalizing sequences—the last two variables in any canalizing sequence can always be reversed. (Why?) □

Exercise 5.15. Find a 3-variable function $f : \mathbb{F}_2^3 \to \mathbb{F}_2$ that satisfies the definition of being nested canalizing from Definition 5.5 except for the last line, requiring $f(x) = \overline{b_3}$. That is,

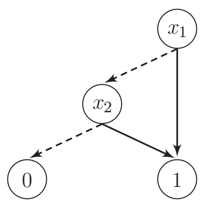
$$f(x) = \begin{cases} b_1 & x_1 = a_1 \\ b_2 & x_1 \neq a_1,\ x_2 = a_2 \\ b_3 & x_1 \neq a_1,\ x_2 \neq a_2,\ x_3 = a_3 \end{cases}$$

What property must f have for this to hold? Draw its BDD with respect to the variable order $x_1 < x_2 < x_3$. Explain why it is reasonable to not call such a function nested canalizing.

Solution. The following function satisfies this:

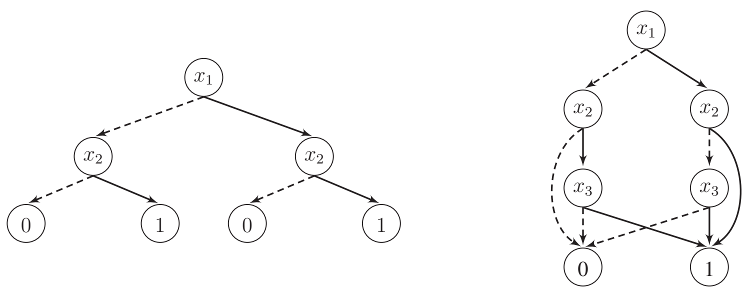
$$f(x) = x_1 \vee x_2 = \begin{cases} 1 & x_1 = 1 \\ 1 & x_1 \neq 1,\ x_2 = 1 \\ 0 & x_1 \neq a_1,\ x_2 \neq a_2,\ x_3 = 1 \end{cases}$$

Inserting another line: 1 $\quad x_1 \neq 1,\ x_2 \neq 1,\ x_3 \neq 1$ would make this the function $x_1 \vee x_2 \vee x_3$. Following is the BDD with respect to the variable order $x_1 < x_2 < x_3$.



□

Exercise 5.16. Determine the Boolean functions described by the following diagrams by writing out the truth table. Are these functions nested canalizing? Simplify the diagrams by merging identical substructures until you wind up with a BDD.

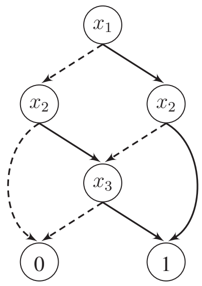


x	**00**	**01**	**10**	**11**
$f(x)$	0	1	0	1

Solution. The first diagram describes the function $f(x_1, x_2) = x_2$. The diagram can be simplified by merging the two x_2-nodes and then eliminating them, but the resulting diagram is not a BDD of a 2-variable function.

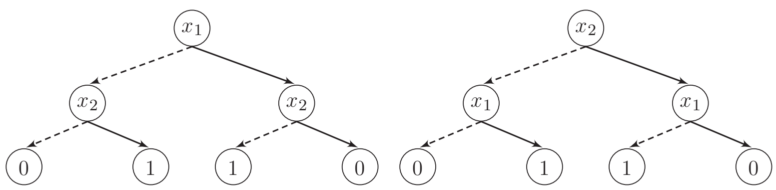
The second diagram describes the Boolean function whose truth table follows: This is the majority function; it is not an NCF. The diagram can be simplified by merging the two x_3-nodes because they have identical substructures.

x	**000**	**001**	**010**	**100**	**011**	**101**	**110**	**111**
$f(x)$	0	0	0	0	1	1	1	1



□

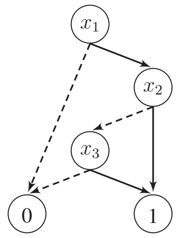
Exercise 5.17. Consider the Boolean function $f(x_1, x_2) = x_1 + x_2$, what is its APL?
Solution. The binary decision trees for $f(x_1, x_2) = x_1 + x_2$ are shown here.

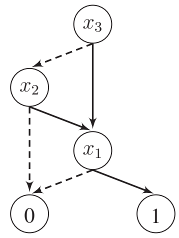


The APL of both is 2, so the $\text{APL}_f = 2$. □

Exercise 5.18. What is the largest APL that a Boolean function of n variables can have? Which functions achieve this?
Solution. The parity function $f(x_1, \ldots, x_n) = x_1 + \cdots + x_n$ has APL n. Every variable must be evaluated before the output can be determined. The complement of the parity function also achieves this. □

Exercise 5.19. Compute the APL of the BDDs shown in Figure 5.8. What is the APL of the Boolean function $f(x_1, x_2, x_3) = x_1x_2x_3 + x_1x_2 + x_1x_3$?
Solution. Following are two BDDs of the function $f(x_1, x_2, x_3) = x_1x_2x_3 + x_1x_2 + x_1x_3$.





The first BDD has APL equals to 7/4. It can be checked that none of the other BDDs (corresponding to the other four variable orders) have a smaller APL, and so $\mathrm{APL}_f = 7/4$. □

Exercise 5.20. For NCFs, any canalizing order will yield a BDD with minimal APL. Can you derive a general formula for the APL of an NCF on n variables?

Solution. Suppose f is an NCF with variable order $x_1 < x_2 < \cdots < x_n$. Then each line in the standard form of the NCF corresponds to a path whose length is given here:

$$f(x) = \begin{cases} b_1 & x_1 = a_1 & \text{path length } 1 \\ b_2 & x_1 \neq a_1,\ x_2 = a_2 & \text{path length } 2 \\ b_3 & x_1 \neq a_1,\ x_2 \neq a_2,\ x_3 = a_3 & \text{path length } 3 \\ \vdots & \vdots & \vdots \\ b_n & x_1 \neq a_1, \ldots,\ x_{n-1} \neq a_{n-1},\ x_n = a_n & \text{path length } n \\ \overline{b_n} & x_1 \neq a_1, \ldots,\ x_n \neq a_n & \text{path length } n \end{cases}$$

The APL of f is the average of all 2^n of these path lengths, which is

$$\mathrm{APL}_f = \frac{1}{2^n}\left(n + \sum_{i=1}^{n} 2^{n-i} i\right) = 2 - \frac{1}{2^{n-1}}$$

□

Exercise 5.21. The NCF $f(x_1, x_2, x_3) = x_1(x_2(x_3 + 1) + 1)$ has only two monomial layers, because of the requirement that the innermost layer must have at least two variables. Find these two layers. That is, write this function as

$$f(x_1, x_2, x_3) = x_1(x_2(x_3 + 1) + 1) = M_1(M_2 + 1) + 1) + b$$

where M_1 and M_2 are extended monomials and $b = 0$ or 1. Find all nested canalizing sequences for f.

Solution. The function f can be written in two extended monomial layers as

$$f(x_1, x_2, x_3) = x_1(x_2(x_3 + 1) + 1) = x_1(x_2\overline{x_3} + 1).$$

That is, $M_1 = x_1$ and $M_2 = x_2\overline{x_3}$, and $b = 0$. There are two nested canalizing sequences: $x_1 < x_2 < x_3$ and $x_1 < x_3 < x_2$. □

Exercise 5.22. Consider the following standard form of an NCF f:

$$f = \begin{cases} 1 & x_1 = 0 \\ 0 & x_1 \neq 0,\ x_2 = 1 \\ 1 & x_1 \neq 0,\ x_2 \neq 1,\ x_3 = 0 \\ 0 & x_1 \neq 0,\ x_2 \neq 1,\ x_3 \neq 0,\ x_4 = 1 \\ 1 & x_1 \neq 0,\ x_2 \neq 1,\ x_3 \neq 0,\ x_4 \neq 1 \end{cases}$$

(a) How many extended monomial layers does f have? How many different nested canalizing sequences does f have in total?

(b) Write down a different nested canalizing sequence of f other than x_1, x_2, x_3, x_4. What is the standard form of f with respect to the sequence you wrote down?

Solution. The extended monomial layers can be read off of the standard form and are

$$f(x_1,\ldots,x_4) = x_1(\overline{x_2}(x_3(\overline{x_4}+1)+1)+1)+1.$$

There are 3 (not 4!) extended monomial layers—the innermost layer, which must have at least two variables, is

$$M_3 = x_3(\overline{x_4}+1) = x_3x_4.$$

There are two nested canalizing sequences: $x_1 < x_2 < x_3 < x_4$, and $x_1 < x_2 < x_4 < x_3$; recall that the last two variables can always be exchanged in any such sequence. To see why, notice that

$$\begin{aligned} x_3(\overline{x_4}+1) &= x_3(x_4+1+1) = x_3x_4 = x_4(x_3+1+1) \\ &= x_4(\overline{x_3}+1). \end{aligned}$$

The standard form with respect to the order $x_1 < x_2 < x_4 < x_3$ is

$$f = \begin{cases} 1 & x_1 = 0 \\ 0 & x_1 \neq 0,\ x_2 = 1 \\ 1 & x_1 \neq 0,\ x_2 \neq 1,\ x_4 = 0 \\ 0 & x_1 \neq 0,\ x_2 \neq 1,\ x_4 \neq 0,\ x_3 = 1 \\ 1 & x_1 \neq 0,\ x_2 \neq 1,\ x_4 \neq 0,\ x_3 \neq 1 \end{cases}$$

The function f, written in layers according to this standard form, is

$$f(x_1,\ldots,x_4) = x_1(\overline{x_2}(x_4(\overline{x_3}+1)+1)+1)+1.$$

□

Exercise 5.23. Compute the partial derivatives $\partial f/\partial x_j$ of each of the following functions. Describe the resulting function; sometimes it will be a commonly known Boolean function.

(i) The logical OR function: $f(x_1,\ldots,x_n) = x_1 \vee \cdots \vee x_n$

(ii) The logical AND function: $f(x_1,\ldots,x_n) = x_1 \wedge \cdots \wedge x_n$

(iii) The parity function: $f(x_1,\ldots,x_n) = x_1 + \cdots + x_n$, where the sum is taken modulo 2

(iv) The *k-threshold* function on n variables, defined by

$$f\colon \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n, \qquad f(x) = \begin{cases} 1 & x_i = 1, \text{ for at least } k \text{ entries of } x \\ 0 & \text{otherwise} \end{cases}$$

All these functions are *symmetric*, which means that $f(x_1,\ldots,x_n) = f(x_{\pi_1},\ldots,x_{\pi_n})$ for all permutations π of $\{1,\ldots,n\}$. Therefore, $\partial f/\partial x_i = \partial f/\partial x_j$ for any i and j.

Solution.

(i) The partial derivative $\partial f/\partial x_1$ of the OR function is

$$\frac{\partial f}{\partial x_1} = \begin{cases} 1 & \text{if } x_2 = x_3 = \cdots = x_n = 0 \\ 0 & \text{else.} \end{cases}$$

This is the function $g(x_2,\ldots,x_n) = \text{nor}(x_2,\ldots,x_n)$.

(ii) The partial derivatives of the AND functions are similar to the OR functions from Part (i), but $= 0$ becomes $= 1$, which means that $g(x_2,\ldots,x_n) = \text{and}(x_2,\ldots,x_n)$.

(iii) Because flipping any bit of an input vector always changes the output of the parity function, its partial derivative $\partial f/\partial x_i = 1$ for all i.

(iv) Flipping the x_1-bit of a k-threshold function changes the output iff exactly $k-1$ other bits are 1. Therefore, each partial derivative $\partial f/\partial x_i$ of the k-threshold function on n-variables is the $(k-1)$-threshold function on $n-1$ variables. □

Exercise 5.24. For each of the following symmetric functions on three variables, compute the activities of a variable (it does not matter which one, due to symmetry) and the compute the sensitivity of the function.

(i) A constant function

(ii) The logical OR function: $f(x_1, x_2, x_3) = x_1 \vee x_2 \vee x_3$

(iii) The logical AND function: $f(x_1, x_2, x_3) = x_1 \wedge x_2 \wedge x_3$

(iv) The parity function: $f(x_1, x_2, x_3) = x_1 + x_2 + x_3 \pmod 2$

(v) The k-threshold function for $k = 1, 2,$ and 3 (See the previous exercise. Note that the 0-threshold and 4-threshold functions are the constant functions.)

Solution.

(i) The activities of each variable and the average sensitivity of a constant function are all zero.

(ii) The activity of x_1 is 1/4 because of all eight vectors in $\mathbb{F}_2^3$, two of them have the property that flipping the x_1 bit flips the output, $\text{or}(x_1, x_2, x_3)$: $(0,0,0)$ and $(1,0,0)$.

For the vector $(0,0,0)$, flipping any of the three bits toggles the output. For each unit basis vector e_1, e_2, e_3, one bit can be flipped to toggle the output. Thus, the average sensitivity is 3/4.

(iii) The activity of x_1 is 1/4, because of all eight vectors in $\mathbb{F}_2^3$, two of them have the property that flipping the x_1 bit flips the output, $\text{or}(x_1, x_2, x_3)$: $(0,1,1)$ and $(1,1,1)$.

For the vector $(1,1,1)$, flipping any of the three bits toggles the output. For each unit basis vector complement $\overline{e_1}, \overline{e_2}, \overline{e_3}$, one bit can be flipped to toggle the output. Thus, the average sensitivity is 3/4.

(iv) The activity of x_1 is 1 because for every vector in $\mathbb{F}_2^3$, flipping any bit flips the output of f. For the same reason, the average sensitivity of f is 3.

(v) The 1-threshold function is OR and the 3-threshold function is AND, so it remains to consider the 2-threshold function. The activity of x_1 is 1/2 because of all eight vectors in $\mathbb{F}_2^3$, four of them have the property that flipping the x_1 bit flips the output: $(0,0,1)$, $(1,0,1)$, $(0,1,0)$, and $(1,1,0)$.

Flipping a bit of $(0,0,0)$ or $(1,1,1)$ will not change the output. For each of the other six vectors, there are two bits that can be toggled to change the output, so the average sensitivity is $12/8 = 3/2$. □

Exercise 5.25. Generalize the previous exercise from 3 to n-variable symmetric functions.

Solution.

(i) The activities of each variable and the average sensitivity of the function are all zero.

(ii) The activity of x_1 is $2/2^n$ because of all 2^n vectors in $\mathbb{F}_2^n$, two of them have the property that flipping the x_1 bit flips $\text{or}(x_1, x_2, x_3)$: $(0,0,\ldots,0)$ and $(1,0,\ldots,0)$.

To compute the average sensitivity, notice that there are n ways to toggle the output by flipping a bit of $(0,0,\ldots,0)$ and one way to toggle the output by flipping a bit for each unit basis vector e_i. No other vectors can be toggled to change the output. Thus, the average sensitivity is $2n/2^n = n/2^{n-1}$.

(iii) The activity of x_1 is $2/2^n$ because of all 2^n vectors in $\mathbb{F}_2^n$, two of them have the property that flipping the x_1 bit flips $\text{or}(x_1, x_2, x_3)$: $(0,1,\ldots,1)$ and $(1,1,\ldots,1)$.

To compute the average sensitivity, notice that there are n ways to toggle the output by flipping a bit of $(1,1,\ldots,1)$, and one way to toggle the output by flipping a bit for each unit basis vector complement, $\overline{e_i}$. No other vectors can be toggled to change the output. Thus, the average sensitivity is $2n/2^n = n/2^{n-1}$.

(iv) The activity of x_1 is 1, because for every vector in $\mathbb{F}_2^n$, flipping any bit flips the output of f. The average sensitivity is n.

(v) In a k-threshold function, flipping the x_1-bit of a vector in $\mathbb{F}_2^n$ flips the output iff exactly $k-1$ other bits are 1. There are $\binom{n-1}{k-1}$ such vectors, and so the activity of x_1 is $2^{-n}\binom{n-1}{k-1}$.

The sensitivity is more complicated. There are two types of vectors for which flipping a bit will change the output:

(a) x has k bits equal to 1; flipping one of these will toggle the output.

(b) x has $k-1$ bits equal to 1; flipping one of the other $n-k+1$ bits equal to 0 will toggle the output.

There are $\binom{n}{k}$ vectors of type (a), and $\binom{n}{k-1}$ vectors of type (b). Thus, the average sensitivity is

$$\begin{aligned}\frac{1}{2^n}\left(k\binom{n}{k}+(n-k+1)\binom{n}{k-1}\right) &= \frac{1}{2^n}\left(\frac{k\cdot n!}{(n-k)!k!}+\frac{(n+k+1)n!}{(k-1)!(n-k+1)!}\right)\\ &= \frac{1}{2^n}\left(\frac{n!}{(n-k)!(k-1)!}+\frac{n!}{(k-1)!(n-k)!}\right)\\ &= \frac{1}{2^{n-1}}\cdot\frac{n!}{(n-k)!(k-1)!}=\frac{k}{2^{n-1}}\binom{n}{k}.\end{aligned}$$

□

Exercise 5.26. Consider the following two NCFs:

$$f_1(x_1,x_2,x_3,x_4)=x_1x_2x_3x_4,\quad f_3(x_1,x_2,x_3)=x_1(x_2(x_3x_4+1)+1),$$

which have one and three extended monomial layers, respectively. Write out the truth table of each function. Then, compute the activity of each variable and the average sensitivity of the function. Make a conjecture as to which layers tend to have the variables with the highest activities. Do you see a trend between the layer number and average sensitivity (and hence stability of Boolean network models that use these functions)? Check Ref. [39] for the answer.

Solution. The function f_1 is just an AND function. Its truth table follows.

x_1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
x_2	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
x_3	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
x_4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
$f(x)$	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**

Flipping 1 bit (does not matter which one) toggles the output exactly 2 of 16 times. The activities are thus

$$\alpha_1^{f_1}=\alpha_2^{f_1}=\alpha_3^{f_1}=\alpha_4^{f_1}=2/16,\quad s^{f_1}=(2+2+2+2)/16=1/2.$$

The truth table of the function f_3 is shown here.

x_1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
x_2	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
x_3	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
x_4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
$f(x)$	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**1**

It is easy to check that changing the x_1-bit toggles the output exactly 10 of 16 times, changing the x_2-bit toggles the output exactly 6 of 16 times, and changing the x_3-bit (or x_4-bit) toggles the output exactly 2 of 16 times. Thus, the activities and sensitivities are

$$\alpha_1^{f_3}=10/16,\quad \alpha_2^{f_3}=6/16,\quad \alpha_3^{f_3}=2/16,\quad \alpha_4^{f_3}=2/16,$$

and the average sensitivity is

$$s^{f_3} = (10 + 6 + 2 + 2)/16 = 20/16 = 5/4.$$

These two examples should give intuition about two key trends about activities and sensitivities of NCFs:

- The activities of the variables in the "outer" layers are higher than those in the "inner" layers;
- The average sensitivity of NCFs with more layers tend to be higher than those with fewer layers. □

Exercise 5.27. Explain why a Derrida curve should always lie below the line $y = x$ for values of ρ close to 1.

Solution. Once almost all of the bits are changed, the resulting vector is so different that it is essentially random, so by dumb luck, roughly half of the output bits should be the same as before the vector was perturbed. In other words, it is very unlikely that changing 95% of the input bits will change 95% of the output bits; it should be much less than that, probably closer to 50%. This means that the Derrida plot will lie below the line $y = x$ for $\rho \approx 1$. □