CS781 Quiz 2 (Autumn 2023)

Max marks: 40 **Duration: 60 mins**

- *The exam is open book and notes. However, you are not allowed to search on the internet or consult others over the internet for your answers.*
- *Be brief, complete and stick to what has been asked.*
- *Unless asked for explicitly, you may cite results/proofs covered in class without reproducing them.*
- ***If you need to make any assumptions, state them clearly.***
- ***Do not copy solutions from others. Penalty for offenders: FR grade.***

1. Consider the neural network shown in Fig. 1. Assume that all hidden layer nodes (x_3 and x_4) have bias 1, the output layer node (x_5) has bias 0, and ReLU activation is used only in hidden layer nodes. As per convention in the Reluplex paper, the inputs and outputs of ReLU are shown separated as x_i^b and x_i^f, for $i \in \{3, 4\}$.

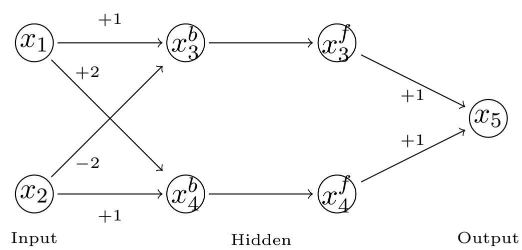

Figure 1: Simple neural network

We wish to check using Reluplex if x_5 can have the value 0, when x_1 and x_2 have values in $[-1, 1]$. You may assume that the values of x_1 and x_2 are set independent of each other.

To help you get going, use the following tableau for (fresh) basic variables b_3, b_4 and b_5.

$$b_3 = x_1 - 2x_2 - x_3^b \qquad b_4 = 2x_1 + x_2 - x_4^b \qquad b_5 = x_3^f + x_4^f - x_5$$

The progress of the Reluplex algorithm can be tracked by filling entries in the following table (using notation from the Reluplex paper), where we have added an extra row ("basic var?") which should have an entry 0 (resp. 1) for a variable iff the variable is non-basic (resp. basic):

variable	x_1	x_2	x_3^b	x_3^f	x_4^b	x_4^f	x_5	b_3	b_4	b_5
basic var?										
lower bound										
assignment										
upper bound										

(a) *[2 marks]* Fill in the table with the right values at the start of Reluplex.

(b) *[3 marks]* Find $\mathsf{slack}^-(b_3)$, $\mathsf{slack}^-(b_4)$ and $\mathsf{slack}^+(b_5)$ at the start of Reluplex.

(c) *[5 marks]* Show the result of applying rule Pivot_2 where b_3 is made non-basic and x_1 is made basic, followed by applying rule Update for b_3. You must show all steps, including the tableau that results after pivoting.

(d) *[2 marks]* Find $\textsf{slack}^+(x_1)$ and $\textsf{slack}^+(b_4)$ at the end of the above sequence of $\textsf{Pivot}_2$ and $\textsf{Update}$.

(e) *[5 marks]* Show the result of applying rule $\textsf{Pivot}_2$ where b_4 is made non-basic and x_2 is made basic, followed by applying rule $\textsf{Update}$ for b_4. You must show all steps, including the tableau that results after pivoting.

(f) *[3 marks]* Indicate (whether by Reluplex or otherwise) whether x_5 can indeed become 0 when x_1 and x_2 vary betwen -1 and 1. Provide justification for your answer.

2. *[10 marks]* Consider the following data for which we wish to construct an optimal binary decision tree. The table below lists values of three binary features f_1, f_2 and f_3 and the binary label l of the corresponding point in the feature space.

f_1	f_2	f_3	l
0	0	0	1
1	1	1	0
1	1	0	1
0	1	1	0

We wish to design an optimal binary decision tree for this data with a single decision node labeled by one of the features, and two leaf nodes labeled by 0 and 1.

Following the ideas in the paper studied in class construct a propositional formula using as few variables as you can, such that a satisfying assignment of the formula gives an optimal decision tree with a single decision node, as desired. You may use sub-formulas of the form $\sum b_i = c$, where each b_i is a propositional variable and c is an integer constant, without expanding the sum into a propositional formula using Sinz' encoding.

You must clearly indicate the meaning of all variables used in your formula.

3. *[10 marks]* Recall the abduction based algorithm studied in class for finding explanations. We wish to apply this algorithm to a complete binary decision tree with two layers of decision nodes that is used to classify the data shown in the table in the previous question. The root node of the tree is labeled f_1 and all children of the root are labeled f_3. Our goal in this question is to find an abduction based explanation of row 3 of the table shown in the previous question (i.e. $f_1 f_2 f_3 = 110$ and $l = 1$).

Find the abductive explanation obtained by applying the algorithm studied in class to the above binary decision tree for row 3. Show all steps clearly.