Proof for Semi-symbolic Node

Kexin Gu Baugh, Nuri Cingillioglu

Say we have a layer of N inputs and each x_i is either 1 or -1. The layer itself:

$$\sum_i^N w_i x_i + \beta = z$$

Conjunction condition: 1. all atom true, the conjunction is true. 2. one atom false, the conjunction is false.

$$\sum_i^N |w_i| + \beta = z \tag{1}$$

$$\sum_{i, i \neq j}^N |w_i| - |w_j| + \beta = -z \tag{2}$$

Equation 1 + 2:

$$\begin{aligned}
-2\beta &= \sum_i^N |w_i| + \sum_{i, i \neq j}^N |w_i| - |w_j| \\
-2\beta &= 2\sum_i^N |w_i| - 2|w_j| \\
\beta &= |w_j| - \sum_i^N |w_i|
\end{aligned}$$

We can use this as a form of β for more general cases.

Say that there are more than one atom in the conjunction being false, this means that the conjunction should still be false. So the layer output should be less than 0. Let $\beta = \alpha - \sum_i^N |w_i|$.

$$\begin{aligned}
\sum_{i \text{ s.t. } w_i x_i > 0}^N |w_i| - \sum_{j \text{ s.t. } w_j x_j < 0}^N |w_j| + \beta &< 0 \\
\sum_{i \text{ s.t. } w_i x_i > 0}^N |w_i| - \sum_{j \text{ s.t. } w_j x_j < 0}^N |w_j| + \alpha - \sum_i^N |w_i| &< 0 \\
\alpha - 2 \sum_{j \text{ s.t. } w_j x_j < 0} |w_j| &< 0 \\
\alpha &< 2 \sum_{j \text{ s.t. } w_j x_j < 0} |w_j| \qquad (3)
\end{aligned}$$

Note that $\alpha - 2\sum_{j \text{ s.t. } w_j x_j < 0} |w_j|$ can be see as the output of a conjunctive layer in general.

Go back to the conjunction condition two, the minimum case would be there is only one atom being false in the conjunction. So Equation 3 becomes:

$$\alpha < 2|w_j| \tag{4}$$

So a suitable value of α for Equation 4 would be $\min_{i,w_i \neq 0} |w_i|$. If $\min_{i,w_i \neq 0} |w_i|$ happen to be the same value of $|w_j|$, we still have inequality $|w_j| < 2|w_j|$. If $w_j = 0$, then we go back to Equation 1 case, and knowing the output of the layer is $\alpha - 2\sum_{j \text{ s.t. } w_j x_j < 0} |w_j|$, we know that the output would still be greater than 0. If we take $\alpha = \min_{i,w_i \neq 0} |w_i|$ and substitute it into Equation 3, the inequality should always hold, as we have verified the base case of only one negation, and adding more to the R.H.S would not change the result of inequality. To conclude, a conjunctive layer's bias should be

$$\beta = \min_{i,w_i \neq 0} |w_i| - \sum_{i}^{N} |w_i| \tag{5}$$

A counterexample for having $\beta = \max_i |w_i| - \sum_i^N |w_i|$ (at least during training time) is like the following. Weights are as 3, 1, 1 and input are as 1, -1, 1.

$$\begin{aligned} & 3 \cdot 1 + 1 \cdot (-1) + 1 \cdot 1 + \beta \\ =& 3 - 1 + 1 + (3 - (3 + 1 + 1)) \\ =& 1 \not< 0 \end{aligned}$$

Although this should be resulting a negative since there is one atom being false, the output layer is greater than 0. However, such problem shouldn't happen if weights are 6 / -6. If the weights are not the same, choice of max and min doesn't matter.