Project 2 for PSYC 489

Bingjun Guo (bingjun3)

February 15, 2024

Part 1

(1)

(a) results are:
$$a_1 = 171.5, a_2 = 171.5, a_3 = 170.5$$

For the first two iteration, $a_1 = 0.5, a_2 = 0.5, a_3 = 1.5$, and then $a_1 = 1.5, a_2 = 1.5, a_3 = 0.5$. Then the smallest input among the 3 nodes will be 2, which is larger than the largest activation, 1.5. Thus, all nodes will keep increasing since their inputs will be constantly larger than their activations.

(b) results are:

$$a_1 = -169.5, a_2 = -170.5, a_3 = -170.5$$

After the first iteration, $a_1 = -0.5, a_2 = 0.5, a_3 = 0.5$; After the second iteration, $a_1 = 0.5, a_2 = -0.5, a_3 = -0.5$; After the third iteratoin, $a_1 = -1.5, a_2 = -0.5, a_3 = -0.5$, and since then, it's clear that sums of inputs and thus activations will drop being negative with an increasing rate because they have been all negative.

(c) results are:

$$a_1 = -340.5, a_2 = -341.0, a_3 = -341.0$$

After the first iteration, a_1 becomes 0 and both a_2 and a_3 are -0.5. Since all nodes are negative, activations will drop with an increasing rate.

(2)

(a) results are:
$$a_1 = 1, a_2 = 1, a_3 = 1$$

Since the activation is bounded above this time, after the two initially activated nodes activate the rest one, all nodes will hold being activated and stays 1.

(b) results are:
$$a_1 = 1, a_2 = 1, a_3 = 1$$

The initially activated node will activate the rest 2, and all 3 stays activated 1 after that.

(c) results are:
$$a_1 = 0, a_2 = 0, a_3 = 0$$

After the first iteration, activations of all units will become 0 since all of them are initially less than 0.75. Activations are bounded below as well, unlike the case in (1).

Part 2

results are:
(a1)
$$a = 0.9999990463256836$$

(a2)
$$a = -56.6650390625$$

(a3)
$$a = 0.9990234375$$

(a4)
$$a = -0.8984375$$

(a5)
$$a = 0.992584228515625$$

(b1)
$$a = -0.9914032911155202$$

(c1)
$$a = -0.81626756145152$$

answers to the questions:

(1)

Activation of $(a1)$ converges to 1 while $(a2)$ becomes negative and gets a large negative value. Because for each iteration of $(a1)$, $\gamma \cdot n_i = 0.75 < 1$, which means that a_1 will keep increasing but never exceed 1 with such input. Meanwhile, for $(a2)$, $\gamma \cdot n_i = 2.5$, and $|(1 - 2.5)| = 1.5 > 1$, which will lead to explosion since the absolute value of Δa will only be larger and larger.

(2)

Activation of $(a3)$ converges to 1 again. $a_1 = \gamma \cdot n_i = 1.5$, and $|(1 - 1.5)| = 0.5 < 1$. With $a_i^t = a_i^{t-1} + 1.5 \cdot (1 - a_i^{t-1}) = 1.5 - 0.5 \cdot a_i^{t-1}$, a_i will get closer to 1, and thus $0.5 \cdot a_i$ will get closer to 0.5, making $1.5 - 0.5 \cdot a_i$ closer to 1. As this logic keeps being consistent, a_i is getting closer to 1.

(3)

I think what happens on $(a5)$ can be considered as gradient vanishing, in which case a_i is too close to 1 and thus $(1 - a_i)$ is too small, making the same input ineffective. The reason why $(a5)$ suffers from this is that the first 10 iteration has already make a_i for $(a5)$ too close to 1, leave the latter inputs ineffective. However, for $(a4)$, input switches to negative after only 2 iterations, which successfully help $(a4)$ escape from the issue.

(4)

Result in $(b1)$ becomes negative instead. In the first 10 iterations of $(b1)$, $\Delta a_i = 0.5 \cdot (1 - a_i) - 0.1 \cdot a_i = 0.5 - 0.6 \cdot a_i$, and thus $a_i^t = 0.5 + 0.4 \cdot a_i^{t-1}$, making a_i slowly increasing towards 1 like it's in $(a5)$. However, after the input of $(b1)$ switches to -1, $\Delta a_i = -0.5 \cdot (1 - a_i) - 0.1 \cdot a_i = 0.4 \cdot a_i - 0.5$, and thus $a_i^t = 1.4 \cdot a_i^{t-1} - 0.5$. Since after only 10 iterations in $(b1)$, $0.4 \cdot a_i < 0.5$, a_i will keep decreasing, become negative, and go on decreasing.

(5)

They will keep decreasing. For $(a4)$, $\Delta a_i = 0.5 \cdot (a_i - 1)$, hence $a_i^t = 1.5 \cdot a_i^{t-1} - 0.5$. For $(b1)$, as is stated previously, $a_i^t = 1.4 \cdot a_i^{t-1} - 0.5$. In both cases, as soon as (a_i) becomes negative, it will go on decreasing faster in exponential manners.

(6)

$(c1)$ will avoid the negative explosion and converge to a value with infinite iterations taking -1 as input. When a_i and input are both negative, $(b1)$ will be unbounded and explode. However, coefficient applied to the input for $(c1)$, will be $(1 + a_i)$ instead, which makes the rule bounded when $-1 < a_i < 0$ and thus avoid the explosion happening to $(b1)$.