Replies to comments – Reviewer 1 (Round 2)

We happily appreciate your great and precious time during the review of our paper and the detailed comments. With respect to the suggestions, we have deeply revised the manuscript. Your comments and questions are printed in the normal font and our replies are in blue print.

Review1.

English language and style
( ) Extensive editing of English language and style required
( ) Moderate English changes required
(x) English language and style are fine/minor spell check required
( ) I don't feel qualified to judge about the English language and style

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	(x)	( )	( )	( )
Is the research design appropriate?	(x)	( )	( )	( )
Are the methods adequately described?	( )	(x)	( )	( )
Are the results clearly presented?	()	(x)	( )	( )
Are the conclusions supported by the results?	( )	(x)	( )	( )

Comments and Suggestions for Authors

The authors have taken into consideration most of my previous remarks. I believe the paper is ready to be accepted. I have only few small/unimportant comments:

There are still no punctuation marks at the end of the highlighted formulas.

Thanks for notifying that punctuation marks must been at the end of formulas. We added punctuation marks to all highlighted formulas.

On line 89: "k/(k-1)/2" most probably should be "k(k-1)/2"

We found that the expression on line 89 was obviously wrong. It is fixed by removing '/'.

One vs One: k(k-1)/2 nodes, True: 1, False: -1, do not classify: 0

In Eq. (9), second line: I prefer the equality sign to be changed to approximation, as used for $R_a(s,s')\approx 0$.

C : The reason why we used approximation instead of equality on the first line was that the actual recursive loop of training algorithm is mostly terminated by the learning rate factor. When π^* (s) is the optimal policy, R (s, s') was considered to be the minimum value although less R (s, s') may be existed in the future. But our mistake was that we overlooked that equation (9) is expressed at the theoretical level. By comment B, we found a ambiguity to equation (9) that the first line was approximate, but the second line marked equality, and determined it have to be equality and corrected as follows.

$$\pi^*(s) = \arg\max_a \{\sum_{s'} P_a(s,s')(R_a(s,s') + \gamma V^*(s'))\} \quad , \text{where } R_a(s,s') = 0 \qquad \textbf{(9)}$$

$$\boldsymbol{\pi^*(s) = \arg\max_a \gamma \sum_{s'} P_a(s,s')\, V^*(s')}$$

On lines 235,236,247: Probably some $ $ marks are forgotten, since the indices of some of the "s" appear in LaTeX style.

These style errors have been corrected.

In the state s_0 of Figure 7, there are two valid actions, one of which is to maintain s_0, and the other is to split its grid and transition to s_0'. In this situation, the reward that can be given is defined as an improvement in accuracy when the Grid Map is transitioned by action.

Eq. (22) represents the Grid Map for s_0.