

This reply is for the reviewer 7.

1. The definition of the transition probability of the product MDP will not be well-defined.
2. The motivation of introducing the augmented LDGBA.
First, the motivation of using LDGBA is to relax the sparsity of rewards. the motivation of our augmentation is to circulate all accepting sets without depending on an MDP or an LTL formula.
3. A formal proof that LDBA and augmented one accepts the same language and the small example of an augmented LDGBA and non-augmented one.
4. The construction augmented LDGBA is similar to the accepting frontier function in [2]. It should be highlighted the difference between the our contribution and the one in [2].

References

- [1] E. M. Hahn, M. Perez, S. Schewe, F. Somenzi, A. Triverdi, and D. Wojtczak, “Omega-regular objective in model-free reinforcement learning,” *Lecture Notes in Computer Science*, no. 11427, pp. 395–412, 2019.
- [2] M. Hasanbeig, Y. Kantaros, A. Abate, D. Kroening, G. J. Pappas, and I. Lee, “Reinforcement learning for temporal logic control synthesis with probabilistic satisfaction guarantee,” *arXiv:1909.05304v1*, 2019.