

This reply is for the reviewer 5.

In the following, we refer to transition-based limit-deterministic generalized Büchi automaton as tLDGBA and non-generalized one as tLDBA.

1. Contributions in our paper.

We propose an augmented tLDGBA. Consequently, our contributions are as follows.

- The sparsity of rewards is relaxed compared to the case of using LDBAs. Therefore, our proposed algorithm is more sample efficient compared to the case of using LDBAs.
- For a product MDP M^\otimes of an MDP M and an augmented tLDGBA \bar{B}_φ corresponding to a given LTL formula and a reward function corresponding to the acceptance condition of M^\otimes , if there exists a finite-memory policy on the MDP M , we obtain a positional policy satisfying φ on M^\otimes by a reinforcement learning algorithm maximizing expected discounted reward.

2. Discussions in the section with the simulation.

If we construct product MDP M^\otimes of an MDP and an LDGBA without any augmentation corresponding to a given LTL formula φ , a policy satisfying φ may not exist depending on M^\otimes and φ . We show that an example in which there is not policies satisfying an LTL formula φ when using the corresponding LDGBA without any augmentation.

3. The motivations of using LDBA.

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, an LTL formula, and a policy.

4. Comparing to the size of state space of the product MDP with the augmented LDGBA and one with the non-augmented LDGBA.

5. Does maximizing the collection of the proposed rewards implies the maximization of the satisfaction probability?

No, it does. It does not generally hold. However, our proposed method can be combined with some research results of attempting to maximize the satisfaction probability such as [1].

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.