For How Long Should We Be Punching? Learning Action Duration in Fighting Games

Authors : Hoang Hai Nguyen, Kurt Driessens and Dennis J.N.J. Soemers

Publication : eprint arXiv
Accepted at Computers and Games 2026

Presenter : s1310115 Oto Ryo

Introduction

- Reinforcement learning is a difficult task for fast-paced games like fighting games.
- Typically, in reinforcement learning for video games, the agent makes a decision about its actions once every N frames.
  ※N is a predetermined positive integer.

If the agent decides its action every frame(N = 1), it gains a high reaction speed, but the computational cost increases.

If N is made larger than 1(frame-skip), the computational cost decreases, but it gains a low reaction speed.

Background

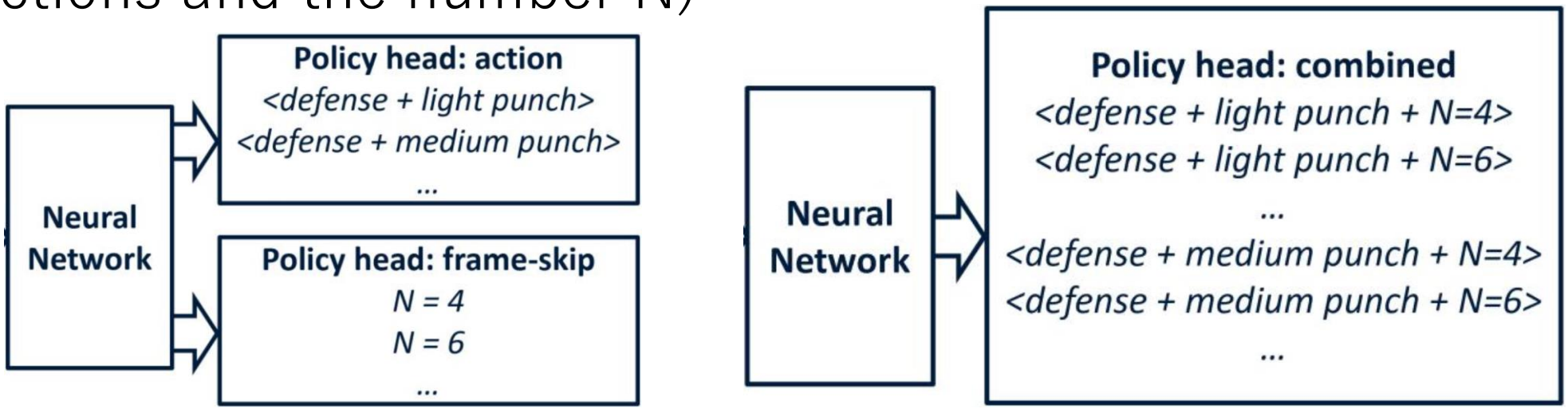
- Reinforcement learning provides a framework for decision-making in which agents interact with the environment in time steps.
- At each time step, the agent observes the state, selects an action, transitions to a new state, and receives a reward based on the state reached.
- The goal is to learn a policy that maximizes the long-term sum of discounted rewards.

Environment

- This research will use Street Fighter 2 as the battle environment and FightLadder for training the AI agent.
- The agent can only observe pixel-based visual state representations.

Experiments

1. Using a fixed frame-skip for a variety of different values for N (4, 8, 16, 60)
2. Using a frame-skip randomly from a predetermined range of possible values for N(4 to 8, 4 to 16)
3. Using a frame-skip from agents decision(agent decides actions and the number N)



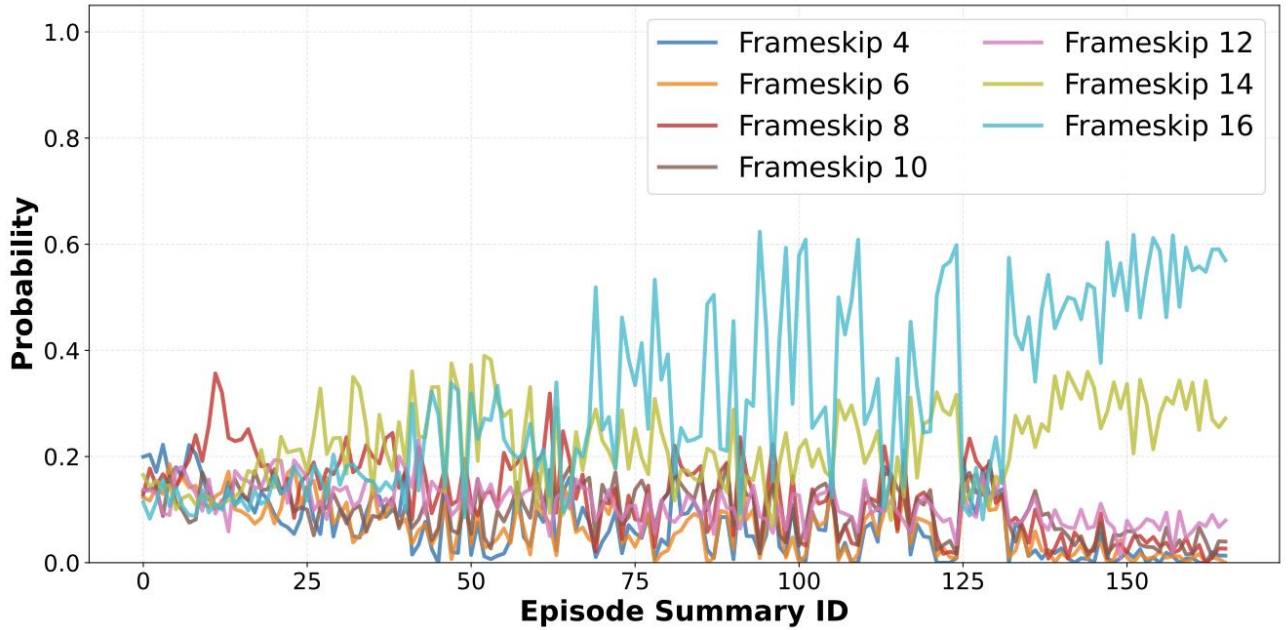
Results

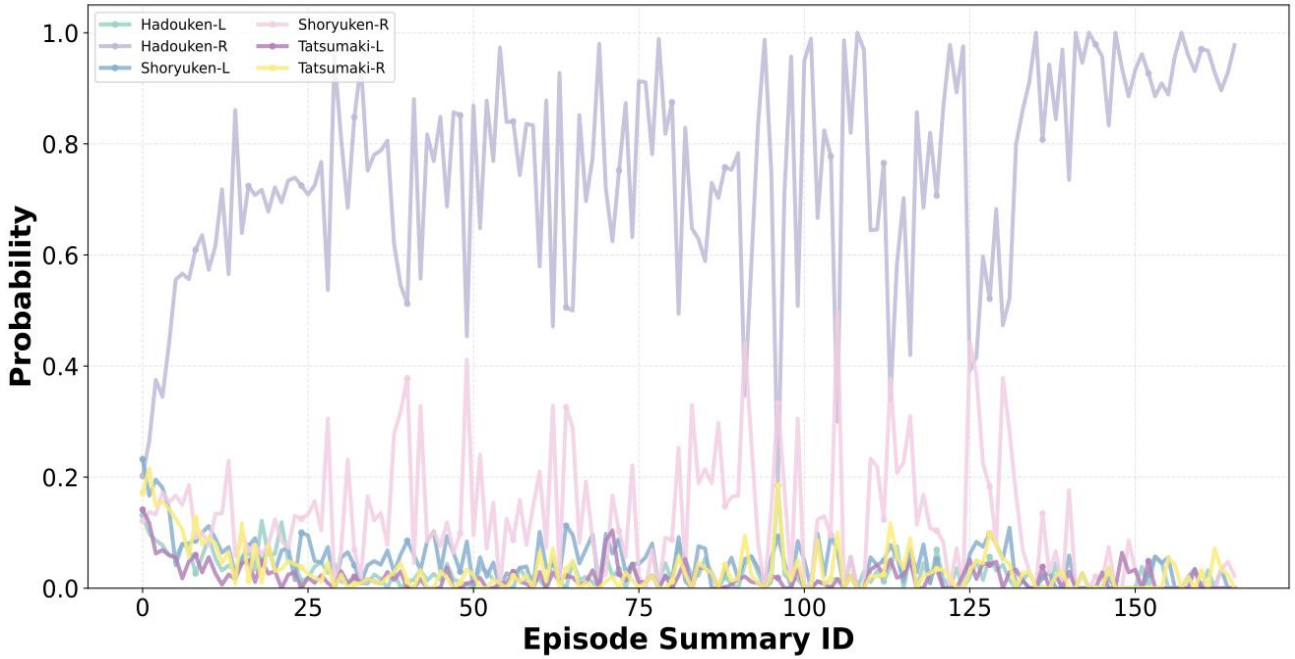
- 95% Agresti-Coull confidence intervals for win percentages of trained agents against Ryu (built-in AI, cycling through star levels 1–8), over 100 evaluation games.

Frame-skip strategy	**Win percentage against Ryu**
Fixed (4)	74% ± 8.5%
Fixed (8)	89% ± 6.4%
Fixed (16)	100% ± 2.6%
Fixed (60)	100% ± 2.6%
Random (4–8)	79% ± 8.0%
Random (4–16)	80% ± 7.9%
Separated (4–8)	77% ± 8.2%
Separated (4–16)	83% ± 7.4%
Separated (4-16,32)	100% ± 2.6%
Combined (4–8)	69% ± 9.0%

Results

- The graph above shows the selected frame-skip for Separated (4 to 16)
- The graph below shows the action selection for Separated (4 to 16).
- Agents repeat the same actions within the game, and there is no variation in their movements or attack types.





Discussion

- The best fixed frame-skip values turn out to be surprisingly high.
- It turns out that even an agent with a frame-skip value of 60 (only making a new decision every full second) obtains a 100% win rate versus the character we train against.
- The agent selected high frame skip values, regardless of hit points or distance to the opponent.

Conclusion

- Reinforcement learning can be used to effectively select frame-skip values autonomously (rather than using predetermined fixed values).
- Agents tended to learn to pick rather high frame-skip values.
- The computer players in this game are weak against exploit strategies that you use over and over.

References

1. Dabney, W., Ostrovski, G., Barreto, A.: Temporally-extended ε-greedy exploration. In: 2021 International Conference on Learning Representations (2021)
2. Kalyanakrishnan, S., Aravindan, S., Bagdawat, V., Bhatt, V., Goka, H., Gupta, A., Krishna, K., Piratla, V.: An analysis of frame-skipping in reinforcement learning. arXiv preprint arXiv:2102.03718 (2021)
3. Li, W., Ding, Z., Karten, S., Jin, C.: Fightladder: A benchmark for competitive multi-agent reinforcement learning. In: Proceedings of the 41st International Conference on Machine Learning. pp. 27653–27674 (2024)
4. McGovern, A., Sutton, R.S.: Macro-actions in reinforcement learning: An empirical analysis. Tech. Rep. 98-70, University of Massachusetts, Amherst (1998)
5. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D.: Human-level control through deep reinforcement learning. Nature **518**, 529–533 (2015)
6. Orenstein, A., Chen, J., Santos, G.A.D., Sapara, B., Bowling, M.: Towards agents that reason about their computation. https://www.arxiv.org/abs/2510.22833 (2025)
7. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 (2017)
8. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of go with deep neural networks and tree search. nature **529**(7587), 484–489 (2016)
9. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al.: A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. Science **362**(6419), 1140–1144 (2018)
10. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, 2 edn. (2018)