AlphaCandy

Mastering the game of Candy Crush without human knowledge

Tim Niklas Witte
wittet@uos.de

Overview

- Already implemented own version of Candy Crush (less candies, see presentation slides for more details).
- Field size: 8x8 $\rightarrow$ 256 ($= 8 \cdot 8 \cdot 4$) actions available.
- Compare different deep reinforcement learing algorithms/network architectures to face a huge action space:
  - Monte carlo tree search (MCTS) (similar to the AlphaGo zero paper [1]).
  - DQN (based on Rainbow [3]):
    - Sequential (consider last 4 time steps):
      - Self-Attention.
      - RNN.
      - Concatenate each time step with each other ($\approx$ feature maps/channels) $\rightarrow$ Convolution.
    - Current state:
      - Self-Attention ("candy-wise attention", similar to DETR [4]).
      - Convolution.
  - Decision Transformer [2] with different desired rewards:
    - x1.5 from best DQN model.
    - x1.5 from MCTS model.
    - x1.5 from best Policy Gradient model.
  - Policy Gradients:
    - TD3.
    - SAC.
    - PPO.

References

[1] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. P. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of go without human knowledge," *Nat.*, vol. 550, no. 7676, pp. 354–359, 2017.

[2] J. D. Brody and D. Perlis, "Decision transformers for glut control in the active logic machine," in *Proceedings of the AAAI 2022 Spring Symposium on Machine Learning and Knowledge Engineering for Hybrid Intelligence (AAAI-MAKE 2022), Stanford University, Palo Alto, California, USA, March 21-23, 2022* (A. Martin, K. Hinkelmann, H. Fill, A. Gerber, D. Lenat, R. Stolle, and F. van Harmelen, eds.), vol. 3121 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022.

[3] M. Hessel, J. Modayil, H. van Hasselt, T. Schaul, G. Ostrovski, W. Dabney, D. Horgan, B. Piot, M. G. Azar, and D. Silver, "Rainbow: Combining improvements in deep reinforcement learning," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018* (S. A. McIlraith and K. Q. Weinberger, eds.), pp. 3215–3222, AAAI Press, 2018.

[4] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part I* (A. Vedaldi, H. Bischof, T. Brox, and J. Frahm, eds.), vol. 12346 of *Lecture Notes in Computer Science*, pp. 213–229, Springer, 2020.