



StARformer: Transformer with State-Action-Reward Representation

Offline RL Imitation Learning Starformer (Ours) DT BC-Vit CQL QR-DQN 1.00 0.75 0.75 0.50 0.50 0.25 0.25 0 00 Atari DMC Atari Performance vs. Seq Length (DMC) Performance vs. Seq. Length (Atari) ▲ Starformer (Ours) ● DT Starformer (Ours) OT 25 20 Applied to real robot Attention Maps (4 heads) State "stav t+3



Results StARformer outperforms DT and other baselines • StARformer successfully scales to longer input sequences Semantic correspondence





Acknowledgements

Intelligence for Improvement of Labor Condition.

References

[1] Chen et al. Decision Transformer: Reinforcement Lear Modeling

[2] Janner et al. Offline Reinforcement Learning as One Big Problem

[3] Touvro et al. Going deeper with image transformers [4] Yang et al. Focal self-attention for local-global interactic

transformers [5] Dosovitskiy et al. An image is worth 16x16 words: Transf recognition at scale







This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Ministry of Science and ICT (No.2018-0-00205, Development of Core Technology of Robot Task-

| ning via Sequence | [6] Agarwal et al. An Optimistic Perspective on Offline Reinforcement Learning |
|---------------------|---|
| g Sequence Modeling | [7] Kumar et al. Conservative Q-learning for Offline Reinforcement Learning |
| | [8] Dabney et al. Distributional Reinforcement Learning with |
| ons in vision | Quantile Regression [9] Kumar et al. Stabilizing Off-policy Q-learning via Bootstrapping |
| formers for image | Error Reduction |
| | |