Weilin YUAN, Jiaxing CHEN, Shaofei CHEN, Dawei FENG, Zhenzhen HU, Peng LI, Weiwei ZHAO, 2024. Transformer in reinforcement learning for decision-making: a survey. *Frontiers of Information Technology & Electronic Engineering*, 25(6):763-790. https://doi.org/10.1631/FITEE.2300548

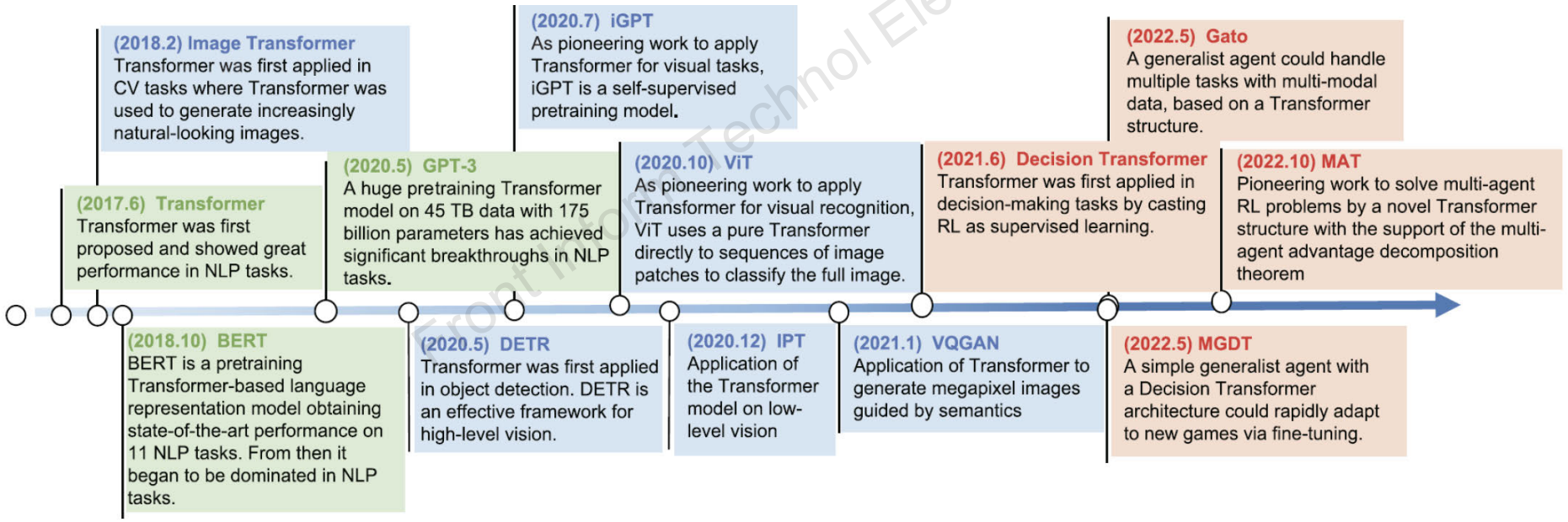
Transformer in reinforcement learning for decision-making: a survey

Key words: Transformer; Reinforcement learning; Decision-making; Deep neural network; Multi-agent reinforcement learning; Meta-reinforcement learning

Weilin Yuan
E-mail: yuanweilin12@nudt.edu.cn
ORCID: https://orcid.org/0000-0001-9894-5253

Transformer development

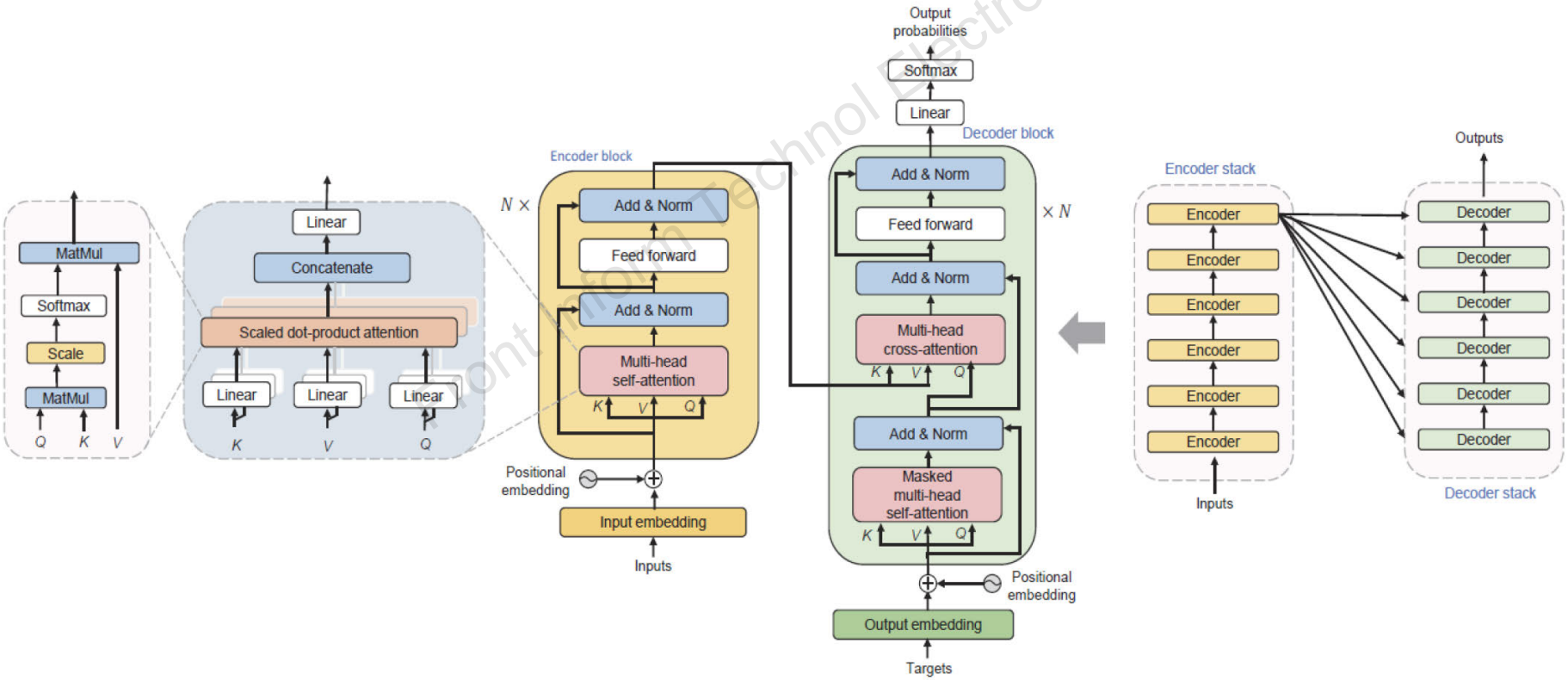
- Inspired by Transformer's major success in the fields of computer vision (CV) and natural language processing (NLP), researchers have made efforts to explore the benefit of Transformer models in solving decision-making (DM) tasks [1-3].



Milestones in the development of Transformer

Transformer structure

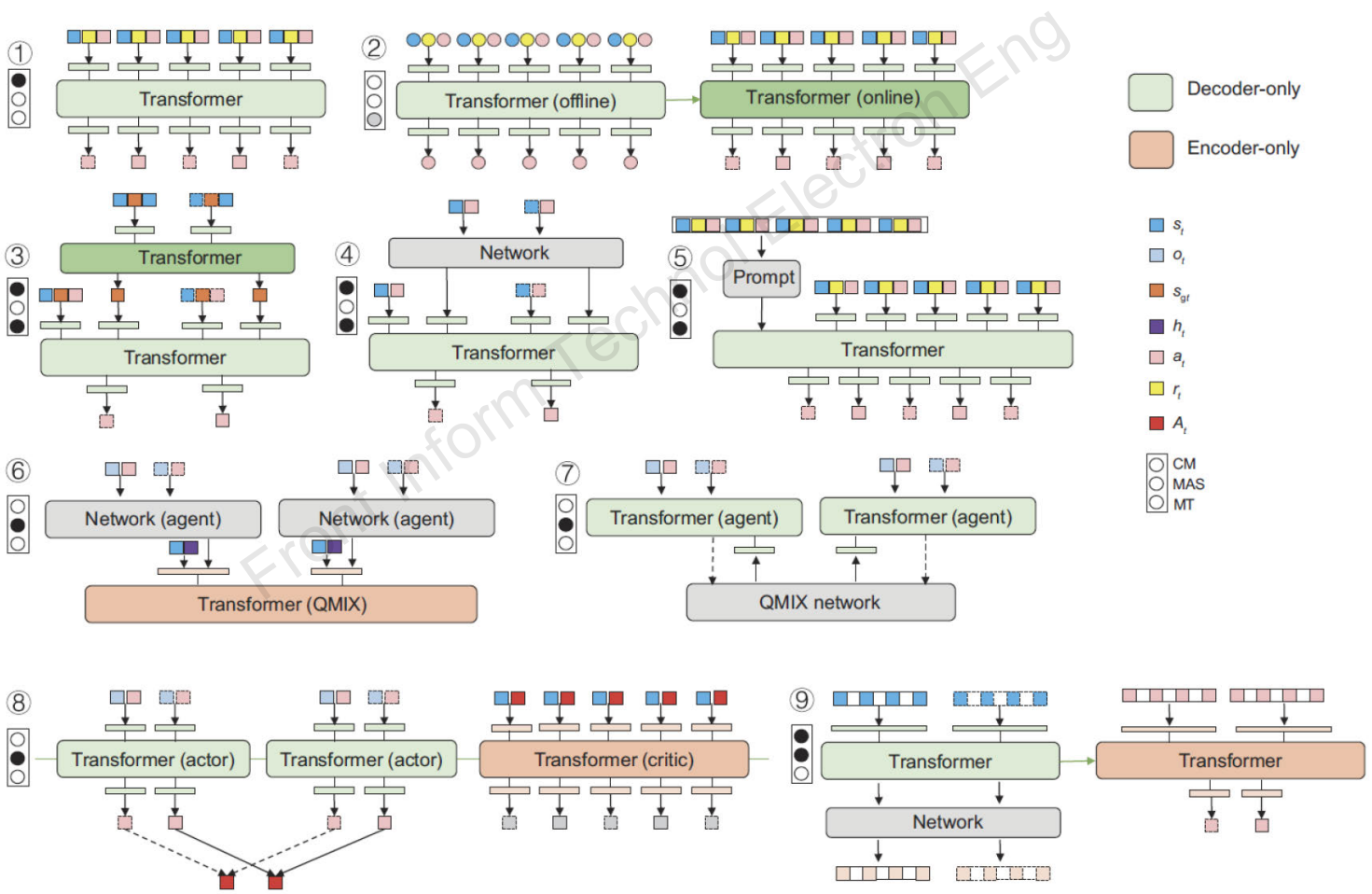
- Transformer [4] is a new kind of neural architecture that sequentially encodes the input sequence as powerful features via the attention mechanism and generates an output sequence after the final encoder output is passed in parallel to a stack of decoders.



Structure of Transformer [4, 5]

Transformer-based RL methods

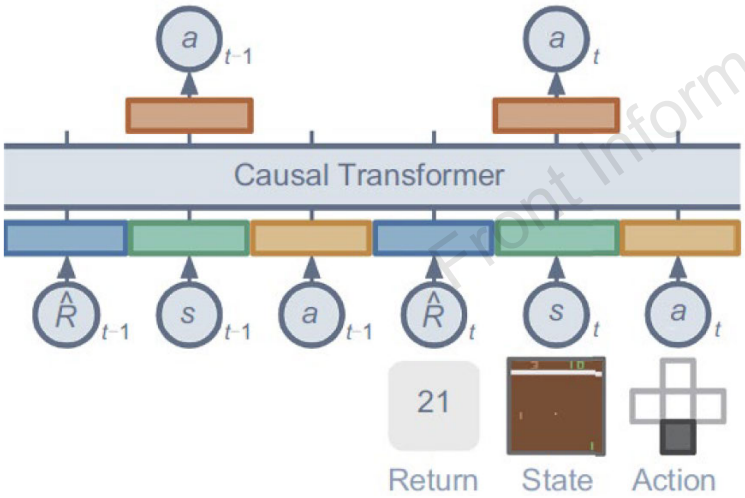
- We provide insights into the advantages and disadvantages of combining Transformer with various RL tasks, such as offline RL, online RL, hierarchical RL, meta-RL, and multi-agent RL (MARL).



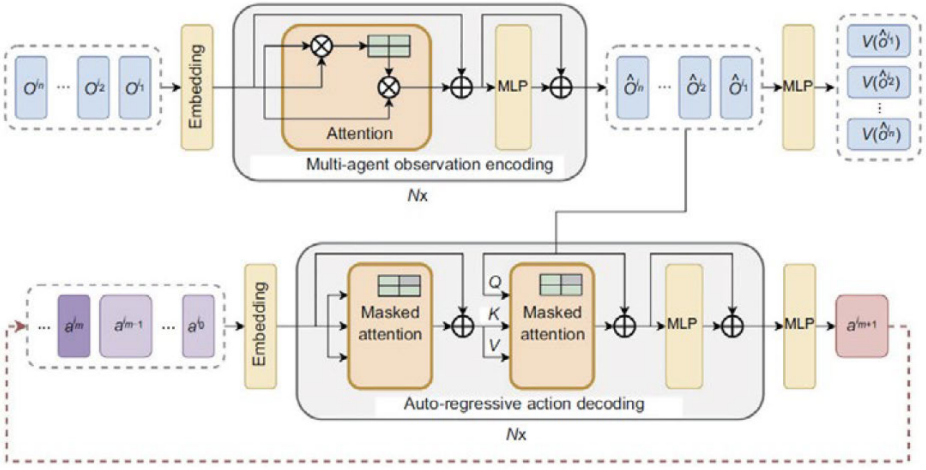
Typical TransRL frameworks

Representative TransRL models

- We provide insights into the details of three representative implementation instances of TransRL models and discuss their characteristics, benefits, and limits.
  - Decision Transformer [6]
  - Multi-agent Transformer [7]
  - Gato [8]



Decision Transformer [6]



Multi-agent Transformer [7]

Applications

- TransRL has been exploited in gaming AI, robotics, transportation, and computer systems.
- We summarize the applications of TransRL in video games, robotic manipulation, robotic navigation, autonomous driving, and combinational optimization.

Domain	RL	TransRL
Gaming AI	• Board games • Card games • Video games	• Video games (open-ended games)
Robotics	• Sim-to-real • Control	• Robotic manipulation • Robotic navigation
Transportation	• Traffic control	• Autonomous driving
Computer systems	• Resource assignment • Security	• Combinatorial optimization

Challenges and open problems

- Stability and structure optimization

->Further improvements

  - Theoretical analysis of TransRL ability
  - Global interaction mechanism design

- Expensive memory and computation

->Further improvements

  - Search-based neural network structure design
  - Automated model compression

- Stochastic effectiveness

->Further improvements

  - TransRL mechanism design for the stochastic environment

We summarize the latest papers and their open-source implementations at https://github.com/williamyuanv0/Transformer-in-Reinforcement-Learning-for-Decision-Making-A-Survey

References

[1] Silver D, Hubert T, Schrittwieser J, et al., 2017. Mastering Chess and Shogi by self-play with a general reinforcement learning algorithm. https://arxiv.org/abs/1712.01815

[2] Jaderberg M, Czarnecki WM, Dunning I, et al., 2019. Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science*, 364(6443):859-865. https://doi.org/10.1126/science.aau6249

[3] Vinyals O, Babuschkin I, Czarnecki WM, et al., 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782):350-354. https://doi.org/10.1038/s41586-019-1724-z

[4] Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

[5] Tay Y, Dehghani M, Bahri D, et al., 2023. Efficient Transformers: a survey. *ACM Comput Surv*, 55(6):109. https://doi.org/10.1145/3530811

[6] Chen LL, Lu K, Rajeswaran A, et al., 2021. Decision Transformer: reinforcement learning via sequence modeling. Proc 34th Int Conf on Neural Information Processing Systems, p.15084-15097.

[7] Wen MN, Kuba JG, Lin RJ, et al., 2022. Multi-agent reinforcement learning is a sequence modeling problem. Proc 36th Int Conf on Neural Information Processing Systems, p.16509-16521.

[8] Reed S, Zolna K, Parisotto E, et al., 2022. A generalist agent. *Trans Mach Learn Res*, 2022:2835-8856.

Weilin YUAN received the MS and PhD degrees in control science and engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2018 and 2023, respectively. He is currently a lecture with NUDT. His current research interests include artificial intelligence, reinforcement learning, opponent modeling, game theory, and information communication.

Weiwei ZHAO received the MS degree from the PLA University of Science and Technology, Nanjing, China, in 2003. He is currently a full professor of the College of Information and Communication with NUDT. His current research interests include artificial intelligence and information communication.