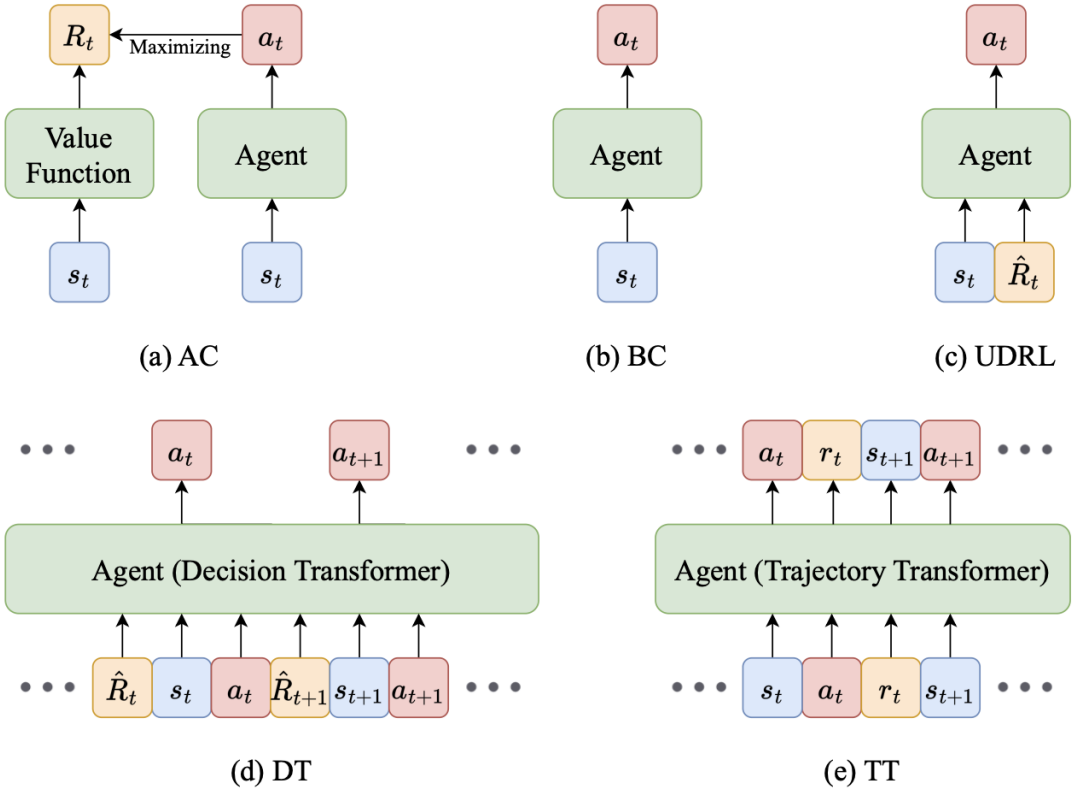
# Large Sequence Models for Sequential Decision-Making: A Survey

**Muning WEN, Runji LIN, Hanjing WANG, Yaodong YANG, Ying WEN, Luo MAI, Jun WANG, Haifeng ZHANG, Weinan ZHANG**

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2689-5

# Problems & Ideas

- Problems:
  - To inquire about the suitability of large sequence models for sequential decision-making and reinforcement learning problems, which are typically beset by long-standing issues involving sample efficiency, credit assignment, and partial observability, etc,.
- Ideas:
  - Presenting a comprehensive overview of recent works aimed at solving sequential decision-making tasks with sequence models such as the Transformer.




Paradigm comparison of conventional RL and a series of works with sequential modeling in the field of sequential decision-making.

# Main Contributions

- Contributions:
  - We discussed the current progress of leveraging the sequence modeling methods for sequential decision-making tasks., and categorizing existing works based on the way they utilize the Transformer.
  - Based on the discussion, we put forth various potential avenues for future research intending to improve the effectiveness of large sequence models for sequential decision-making, encompassing theoretical foundations, network architectures, algorithms, and efficient training systems.
  - We hope this survey could inspire more investigation into this trending topic and ultimately empower more real-world applications.

| Method | Sequence | Prediction | Discretized Tokens | Benefit | Notes |
|---|---|---|---|---|---|
| UPDeT [52] | s | a | No | Multi-task; Few-shot learning; Interpretability | Model-free; Online; Multi-agent |
| PIT [53] | s | Q values | No | Multi-task; Few-shot learning; Credit assignment | Model-free; Online; Multi-agent |
| DT [44] | rtg-s-a | a | No | Long sequence; POMDP; Credit assignment | Model-free; Offline |
| TT [45] | s-a-r(-rtg) | s-a-r | Yes | Long sequence; POMDP; Sparse-reward | Model-based; Offline |
| GDT [59] | $\psi(s,a)$-s-a | a | No | HIM problems | Model-free; Offline |
| PDT [46] | s-a | a | No | Few-shot learning | Model-free; Pre-train |
| MADT [50] | s-a | a | No | Multi-task; Long Sequence | Model-free; Offline; Multi-agent |
| ODT [49] | rtg-s-a | a | No | Few-shot learning | Model-free; Online |
| MAT [54] | s | a | No | Monotonic improvement; Multi-Task; Few-shot learning | Model-free; Online; Multi-agent |
| MGDT [55] | s-a-r-rtg | a-r-rtg | Yes | Multi-task; Few-shot learning | Model-free; Offline |
| TrMRL [60] | s | a | No | Multi-task; Few-shot learning | Model-free; Online; Meta-learning |
| PG-AR [61] | s | a | No | Monotonic improvement | Model-free; Online; Multi-agent |
| Prompt-DT [56] | rtg-s-a | a | No | Multi-task; Few-shot learning | Model-free; Offline |
| BooT [62] | s-a-r-rtg | s-a-r-rtg | Yes | Data Augmentation | Model-based; Offline |