Figure 1: Visual demonstration and comparison for the credit assignment process behind different methods.
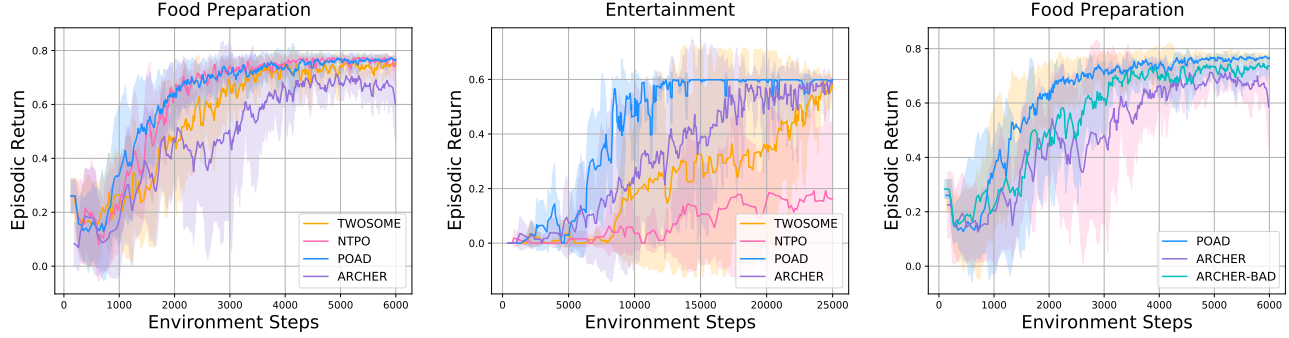

Figure 2: Adding baseline ArCHer on VirtualHome (left two), and the extension of integrating ArCHer with BAD (right one).
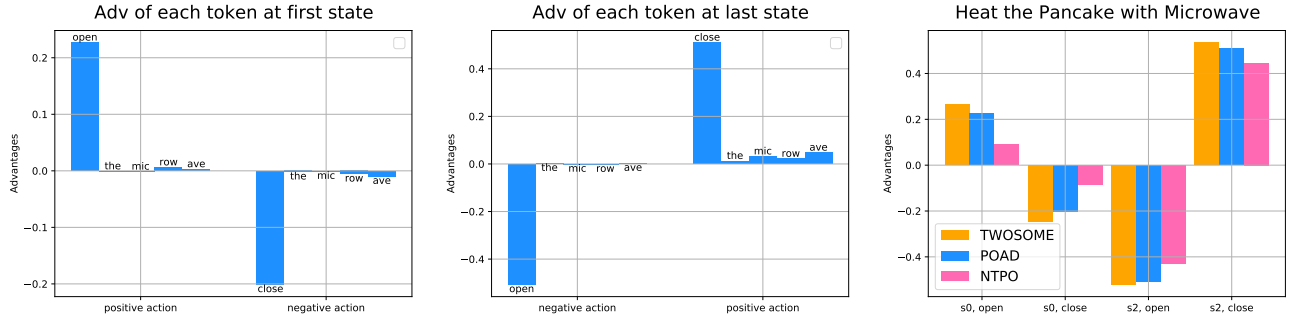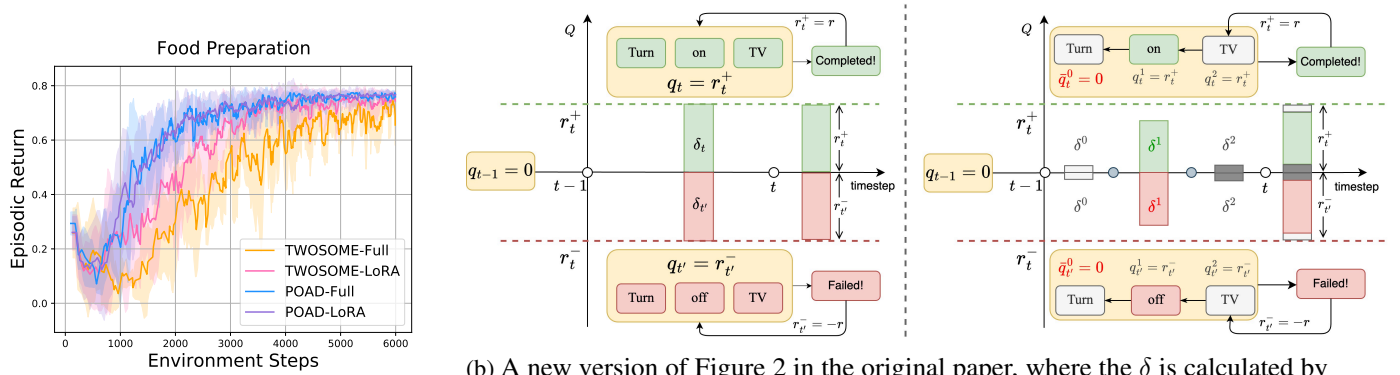

Figure 3: Case Study: Demonstration of the token-level credit assignment learned by the BAD at two states (left two). And a comparison of the volume of credit assignment for key tokens between different methods (right one), where TWOSOME indicates the credit assigned to entire actions instead of specific tokens.

(a) Ablation on full-scale fine-tuning and LoRA fine-tuning.

(b) A new version of Figure 2 in the original paper, where the δ is calculated by subtracting the q value of the previous (left) action/token from the q value of the current action/token.

Figure 4