

Table 1: Real-world Evaluation

Language-conditioned						Goal-conditioned				
Task	TRA	GRIF	LCBC	Octo	AWR	TRA	GRIF	GCBC	Octo	AWR
<b>(A)</b> open the drawer	<b>0.80</b> <sup>(±0.1)</sup> †	0.20 <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup> †	0.60 <sup>(±0.2)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	<b>0.50</b> <sup>(±0.2)</sup>	<b>0.80</b> <sup>(±0.2)</sup>
<b>(A)</b> mushroom in drawer	<b>0.80</b> <sup>(±0.1)</sup>	<b>0.80</b> <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.90</b> <sup>(±0.1)</sup>	0.40 <sup>(±0.2)</sup>	<b>0.80</b> <sup>(±0.2)</sup>	<b>0.90</b> <sup>(±0.1)</sup>	<b>0.60</b> <sup>(±0.2)</sup>
<b>(A)</b> close drawer	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	<b>1.00</b> <sup>(±0.0)</sup>	0.40 <sup>(±0.2)</sup>	0.80 <sup>(±0.2)</sup>	0.60 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>
<b>(D)</b> take the item out of the drawer	<b>0.60</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>
<b>(B)</b> put the spoons on towels	<b>1.00</b> <sup>(±0.0)</sup>	0.40 <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	<b>1.00</b> <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.60 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.60 <sup>(±0.2)</sup>
<b>(B)</b> put the spoons on the plates	<b>0.80</b> <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>1.00</b> <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.40 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.80 <sup>(±0.2)</sup>
<b>(C)</b> put the corn and sushi on plate	<b>0.90</b> <sup>(±0.1)</sup>	0.00 <sup>(±0.0)</sup>	0.40 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.50 <sup>(±0.2)</sup>	<b>0.70</b> <sup>(±0.1)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.30 <sup>(±0.1)</sup>
<b>(C)</b> sushi and mushroom in bowl	<b>0.80</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>
<b>(C)</b> corn, banana, and sushi in bowl	<b>0.80</b> <sup>(±0.1)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.1)</sup>	<b>0.50</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	0.50 <sup>(±0.2)</sup>
<b>(D)</b> corn on plate then sushi in pot	<b>0.70</b> <sup>(±0.1)</sup>	0.00 <sup>(±0.0)</sup>	0.40 <sup>(±0.2)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	<b>0.30</b> <sup>(±0.1)</sup>	<b>0.20</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>
<b>(A)</b> sweep to the right	<b>0.80</b> <sup>(±0.1)</sup>	0.20 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.70</b> <sup>(±0.1)</sup>	0.40 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.80</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>
<b>(B)</b> fold cloth into the center	<b>1.00</b> <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	0.40 <sup>(±0.2)</sup>	<b>1.00</b> <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.60 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>
<b>(B)</b> move bell pepper and sweep towel	<b>0.50</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	0.00 <sup>(±0.0)</sup>	0.20 <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>	<b>0.60</b> <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	0.20 <sup>(±0.2)</sup>	<b>0.40</b> <sup>(±0.2)</sup>	0.00 <sup>(±0.0)</sup>

**(A)** One step tasks  
**(C)** Semantic generalization

**(B)** Task concatenation  
**(D)** Tasks with dependency

†The best-performing method(s) up to statistical significance are **highlighted**