

A COMPUTATIONAL RESOURCE USAGE AND REPRODUCIBILITY REQUIREMENTS

Experiments were run on our internal GPU cluster. Running a meta-learning experiment to 30,000 iterations takes about 3 days on a NVIDIA Tesla V100 GPU. For 7 different experiment runs with 10 seeds each, the total compute time is about 210 GPU-days, though in practice the experiments can be run in parallel.

The batch size (4096) we use is quite large and does not fit in GPU memory for a consumer grade GPU. In order to achieve these batch sizes, we use gradient accumulation, so the batch size for each backward step might be 256, but then gradients are averaged over 16 steps to make an effective batch size of 4096. This trades training time for optimization stability.

B ADDITIONAL COMPARISONS

We show additional related work comparisons and hyperparameters in Tables 4 and 5.

	seq2seq 2020	GECA 2020	FILM 2021	RelNet 2021	LCGN 2020	Planning 2020	RD Random/RL 2022	Ours(o) Ours
Learning Rate	0.001	0.001	0.0015	0.0015	0.0001	0.001		0.0001
Batch Size	200	200	128	128	64	200	512	4096
Steps	200K	200K	115K	115K	575K	368K	13.2M	35K
#params						535K		
A	97.15 \pm 0.46	87.6 \pm 1.19	98.83 \pm 0.32	97.38 \pm 0.33	98.6 \pm 0.9	94.19 \pm 0.71	98.39 \pm 0.17	0.95 \pm 0.01
B	30.05 \pm 26.76	34.92 \pm 39.30	94.04 \pm 7.41	49.44 \pm 8.19	99.08 \pm 0.69	87.31 \pm 4.38	62.19 \pm 24.08	0.96 \pm 0.01
C	29.79 \pm 17.70	78.77 \pm 6.63	60.12 \pm 8.81	19.92 \pm 9.84	80.31 \pm 24.51	81.07 \pm 10.12	56.52 \pm 29.70	0.97 \pm 0.01
D	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00	0.00 \pm 0.00	0.16 \pm 0.12		43.60 \pm 6.05	0.35 \pm 0.05
E	37.25 \pm 2.85	33.19 \pm 3.69	31.64 \pm 1.04	42.17 \pm 6.22	87.32 \pm 27.38	52.8 \pm 9.96	53.89 \pm 5.39	0.85 \pm 0.20
F	94.16 \pm 1.25	85.99 \pm 0.85	86.45 \pm 6.67	96.59 \pm 0.94	99.33 \pm 0.46		95.74 \pm 0.75	0.97 \pm 0.01
H	19.04 \pm 4.08	11.83 \pm 0.31	11.71 \pm 2.34	18.26 \pm 1.24	33.6 \pm 20.81		21.95 \pm 0.03	0.86 \pm 0.02

Table 4: Additional related work comparisons.

	ViT Qiu et al., 2021	Modular Ruis & Lake, 2022	Role-guided Kuo et al., 2021	Transformer Ours	Ours(o, A) Ours	Ours(o) Ours
Learning Rate	0.0015	0.001	0.001	0.0001	0.0001	0.0001
Batch Size	128	200	200	4096	4096	4096
Steps	114.96K	73K	150K	35K	35K	35K
#params	3M			13.2M	13.2M	13.2M

Table 5: Hyperparameters used in the related work comparisons of Table 2.

C ORACLE FUNCTION

The oracle function generates relevant instructions by the use of a templating mechanism, which replaces verbs and adverbs in the sentence with other verbs and adverbs, such that the whole combination is still in distribution, but not the same as the query instruction. The rules of the system are:

- Replace “pull” with “push” and “walk to”
- Replace “walk to” with “push” and “pull” (but not if “while spinning” is the adverb)
- Replace “push” with “walk to” and “pull” (but not if “while spinning” is the adverb)
- Replace “while zigzagging” with “hesitantly”, nothing and “while spinning” (but not if “push” is the verb)
- Replace “hesitantly” with “while zigzagging”, nothing and “while spinning” (but not if “push” is the verb)
- Replace “while spinning” with “hesitantly”, “while zigzagging” and nothing

Examples of what the oracle function generates for a given query instruction and environment can be found in Table 6.

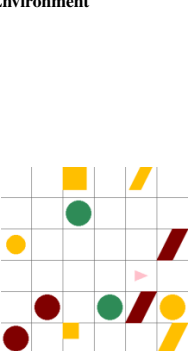
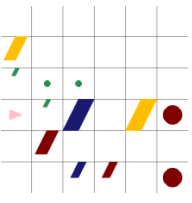
Environment	Query Instruction	Target Actions	Supports
	walk to a small circle hesitantly	LTURN(2) WALK STAY WALK STAY WALK STAY WALK STAY RTURN WALK STAY	push a small circle hesitantly LTURN LTURN WALK STAY WALK STAY WALK STAY WALK STAY RTURN WALK STAY PUSH STAY PUSH STAY PUSH STAY PUSH STAY pull a small circle hesitantly LTURN LTURN WALK STAY WALK STAY WALK STAY WALK STAY RTURN WALK STAY pull STAY pull STAY pull STAY pull STAY walk to a small circle while spinning LTURN(4) LTURN LTURN WALK LTURN(4) WALK LTURN(4) WALK LTURN(4) WALK LTURN(4) RTURN WALK walk to a small circle while zigzagging LTURN LTURN WALK RTURN WALK LTURN WALK WALK WALK LTURN LTURN WALK walk to a small circle WALK WALK WALK RTURN WALK
	push a red small cylinder while zigzagging	WALK RTURN WALK LTURN WALK RTURN WALK LTURN WALK PUSH	walk to a red small cylinder while zigzagging WALK RTURN WALK LTURN WALK RTURN WALK LTURN WALK push a red small cylinder while spinning LTURN(4) WALK LTURN(4) WALK LTURN(4) WALK LTURN(4) RTURN WALK LTURN(4) WALK push a red small cylinder hesitantly WALK STAY WALK STAY WALK STAY RTURN WALK STAY WALK STAY push a red small cylinder WALK WALK WALK RTURN WALK WALK

Table 6: Examples of what supports the oracle function generates for a given query instruction and environment state. These two examples are from the training data. Note that we never generate the same instruction as the query instruction in the supports, and we also never generate any Split H instruction in the supports. Also note that in some cases, the environment makes pushing or pulling an object impossible, even though it is in the instruction, see the second row for an example of this.