Neuro-symbolic Learning of Lifted Action Models from Visual Traces
Supplementary Material

Primary Keywords: *(2) Learning;*

Ablation Study

We conducted an ablation study on various components of ROSAME-I to justify our design choices. Specifically, we aim to demonstrate the impact of the additional prior bias introduced in the paper and the effects of overparameterizing PAMs with neural networks instead of directly learning the probability distributions over the four PAM cases for each action symbol.

Prior Bias

We demonstrate the effect of the additional prior bias by training another ROSAME-I with $\lambda = 0$ (without prior bias loss), keeping all other settings the same as in the original experiments. We then compare the results obtained in this scenario with the evaluation results presented in the paper, both of which are included in Tab. 1.

As we can see, the action models learned without prior bias loss contain more errors, even for the Blockworld domain where there is no prevail condition. The result suggests that the prior bias not only assists in prevail condition learning but also enhances the overall quality of the recovered model.

	Prior Bias Setting	Acc	Error
Blockworld (grid world)	λ**=0.2**	**97.51%**	**0**
	λ=0	94.36%	1
Gripper	λ**=0.2**	**90.54%**	**0**
	λ=0	89.73%	3
Logistics	λ**=0.2**	**96.41%**	**0**
	λ=0	94.07%	10
Tower of Hanoi	λ=0.2	99.72%	**1**
	λ=0	**99.98%**	3
8-puzzles	λ**=0.4**	**99.67%**	**4**
	λ=0	98.58%	20

Table 1: Comparison of ROSAME-I's performance between training with and without prior bias.

We also observe that the CV model's prediction accuracy generally improves with the inclusion of the prior bias loss, except for the Tower of Hanoi domain. This is unsurprising because the quality of the evolving model in ROSAME directly affects the CV model's learning.

In Tab. 2, we compare the two models learned for the Tower of Hanoi domain in this experiment. By reducing λ to 0, we effectively recover a model indistinguishable from the model on the left where $\lambda = 0.2$, with all the prevail conditions recognised as add effects. This comparison explains why the accuracy of the CV model does not drop, as both models can explain any valid traces in this domain.

Overparameterization

We implement a PAM model that directly learns the four values of the discrete probability distribution over the four PAM cases for each action symbol. Compared to the PAM network presented in the paper, this model maintains a learnable vector of length four.

We train ROSAME-I with this PAM model instead of the PAM network for each domain, as shown in Tab. 3. We observe that ROSAME-I, when using this new PAM model, achieves significantly lower state prediction accuracy and recovers action models of lower quality. Overparameterizing PAM into a PAM network facilitates our learning process and is crucial for ROSAME-I to converge to high-quality solutions.

During this ablation study, we also find that sometimes (but not always), when ROSAME-I with the new PAM model gets stuck, re-initializing ROSAME (i.e., the new PAM model but not the CV model) may help ROSAME-I overcome the local optimum and converge to a better solution. By repeating this re-initialization process several times during training, we can bring the new PAM model's performance closer to that of the PAM network. This phenomenon further confirms our hypothesis on overparameterization. Re-initializing ROSAME encourages PAM to explore a different model while keeping the CV model unchanged helps exploit the knowledge already acquired in the previous training.

Reasoning Shortcuts and Data Augmentation

Reasoning shortcuts are a specific issue within neuro-symbolic methods (Li et al. 2023; Marconato et al. 2023). These shortcuts occur when a deep learning model incorrectly maps inputs to high-level concepts, yet the overall model still achieves low loss and high accuracy on training data by combining incorrect concepts with incorrect symbolic inference. In our specific context, reasoning shortcuts occur when the CV model makes incorrect state predictions and ROSAME learns a wrong action model, in which case the two components work together to produce a consistent

$\lambda = 0.2$	$\lambda = 0$
move(?a ?b ?c) precondition: **(smaller ?b ?a) (smaller ?c ?a)** (on ?a ?b) **(clear ?a)** (clear ?c) add effect: (clear ?b) (on ?a ?c) delete effect: (on ?a ?b) (clear ?c)	**move(?a ?b ?c)** precondition: (on ?a ?b) (clear ?c) add effect: (clear ?b) (on ?a ?c) **(clear ?a)** **(smaller ?b ?a) (smaller ?c ?a)** delete effect: (on ?a ?b) (clear ?c)

Table 2: Comparison between the two models learned with different λ in the Tower of Hanoi domain.

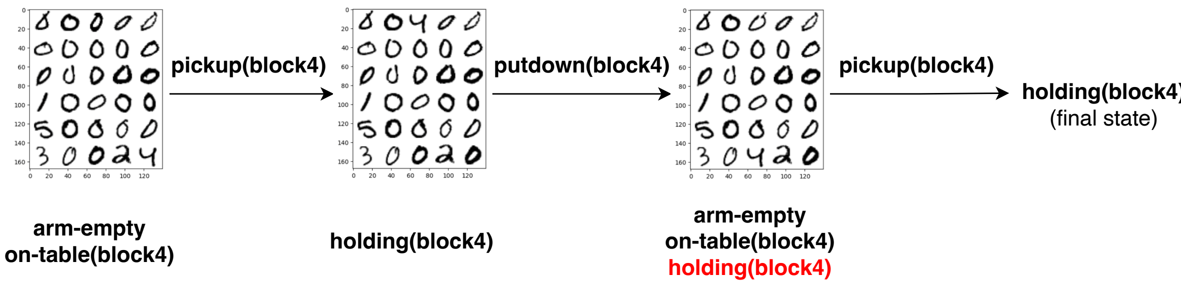

Figure 1: A 3-step trace and partial proposition predictions corresponding to a Reasoning Shortcut action model in the Blockworld domain.

	Overparameterization	Acc	Error
Blockworld (grid world)	PAM model	82.83%	20
	PAM network	**97.51%**	**0**
Gripper	PAM model	76.95%	9
	PAM network	**90.54%**	**0**
Logistics	PAM model	92.79%	12
	PAM network	**96.41%**	**0**

Table 3: Comparison of ROSAME-I's performance between using the new PAM model and the PAM network

but unreasonable trace, leading to the correct goal state.

Fig. 1 gives an example of a 3-step trace corresponding to a learned reasoning shortcut model in the Blockworld domain, where the $\mathrm{putdown}(?block)$ action schema misses a delete effect: $\mathrm{holding}(?block)$. As the predictions are consistent with the incorrect action model learned by ROSAME, and the goal state is correctly predicted, no loss is incurred. ROSAME-I becomes stuck around this sub-optimal solution. Leaving reasoning shortcuts during training is relatively challenging because if ROSAME adjusts its action model, the predictions and the model will become inconsistent. ROSAME-I usually needs to incur higher losses before moving towards the globally optimal solution. In some neuro-symbolic tasks, the reasoning short model(s) may even achieve global optimal loss as well (hence indistinguishable from the ground truth solutions) (Marconato, Teso, and Passerini 2023).

We discover that one specific reason for reasoning shortcuts in our task is the CV model's failure to generalize among different images representing the same state. If we look at the first and the third images in Fig. 1, we realize that they represent the same state, but the images are different due to the change of $block4$'s position when it is placed on the ground by the arm. However, our CV model fails to draw the connection between the two images, leading to inconsistent predictions for these images to match the problematic action model in ROSAME. If the CV model had better generalization, it would identify this discrepancy, and ROSAME-I would not settle for this reasoning shortcut solution. This type of problem is not uncommon in our experiments, especially for the grid world representation where the grids are discrete. To mitigate the issue of reasoning shortcuts and enhance the CV model's generalization, we create customized data augmentation methods for domains using grid world representations. In the Blockworld domain, we randomly alter the positions and order of the block towers. In the Gripper domain, we randomly change the balls' positions in each room. In the Logistics domain, we randomly rearrange the positions of all items within each 3×3 grid associated with each location.

References

Li, Z.; Liu, Z.; Yao, Y.; Xu, J.; Chen, T.; Ma, X.; and Lü, J. 2023. Learning with Logical Constraints but without Shortcut Satisfaction. In *Proc. ICLR*. OpenReview.net.

Marconato, E.; Bontempo, G.; Ficarra, E.; Calderara, S.; Passerini, A.; and Teso, S. 2023. Neuro-Symbolic Continual Learning: Knowledge, Reasoning Shortcuts and Concept Rehearsal. In *Proc. ICML*, volume 202 of *Proceedings of Machine Learning Research*, 23915–23936. PMLR.

Marconato, E.; Teso, S.; and Passerini, A. 2023. Neuro-Symbolic Reasoning Shortcuts: Mitigation Strategies and their Limitations. In *Proc. NeSy*, volume 3432 of *CEUR Workshop Proceedings*, 162–166. CEUR-WS.org.