ZERO-SHOT SYNTHETIC POLYRULE REASONING WITH NEURAL-SYMBOLIC INTEGRATION

Anonymous authors
Paper under double-blind review

ABSTRACT

We propose a neural-symbolic framework for zero-shot learning in Synthetic PolyRule Reasoning (SPR). Our approach integrates neural networks for feature extraction with a symbolic reasoning component to infer and apply unseen rules without retraining, thus tackling new tasks in a zero-shot manner. We demonstrate its potential on the SPR_BENCH dataset by focusing on a single metric (Shape-Weighted Accuracy or Color-Weighted Accuracy) and show how the model can adapt to tasks governed by novel symbolic constraints.

1 INTRODUCTION

Recent advances in language modeling have drawn attention to zero-shot reasoning, where systems are expected to handle new tasks without additional training (**?**). Zero-shot capabilities are particularly relevant for real-world deployment, since data collection for every possible rule can be costly. However, purely neural systems often struggle with novel or complex rule structures, leading to pitfalls such as overfitting to seen rules and failing to generalize (**?**).

Synthetic PolyRule Reasoning (SPR) is a domain where shape-color sequences must be classified according to various symbolic constraints. Errors emerge in real-world-inspired scenarios when new constraints deviate significantly from prior training data. We explore a neural-symbolic integration strategy that aims to handle these challenges. Our findings highlight both the promise and pitfalls of combining neural representations with explicit rule engines.

Contributions We present the first exploration of zero-shot SPR using a neural-symbolic approach. Our key findings include: (1) a method for incorporating symbolic rule instantiation (**?**) into a neural model for classification, (2) an empirical demonstration that the system can handle previously unseen rules, and (3) practical insights into pitfalls that arise from insufficient rule coverage in training and from incremental modifications of symbolic constraints.

2 RELATED WORK

Zero-shot reasoning has been tackled in purely neural approaches (**?**) and in hybrid settings (**??**). Recent neuro-symbolic methods focus on compositional tasks, combining learned representations with explicit rule handling (**??**). Other efforts incorporate symbolic structures to ensure interpretability and flexible integration of constraints (**??**). Our work applies these ideas to SPR, highlighting both partial successes and pitfalls in zero-shot adaptation.

3 METHOD

Our method combines (1) a neural feature extractor for shape-color pairs and (2) a symbolic inference engine for rule instantiation (**?**). The neural backbone produces embeddings for each pair, while the symbolic engine matches a set of rule templates to these embeddings. Zero-shot generalization arises from applying new rule templates beyond those observed during training.

The key to avoiding pitfalls lies in maintaining rule representations that are sufficiently generic, so that minor extensions or modifications do not break the pipeline. However, when constraints vastly differ from those in training data, or when a completely new shape-color binding is introduced, we found that the model suffers from a sharp performance drop.

4 EXPERIMENTS

We evaluate on SPR_BENCH (**?**), which includes train, dev, and test splits with multiple symbolic constraints. Training uses a partial set of rules; the evaluation includes unseen rules. The main metric is a shape-weighted or color-weighted accuracy that captures how well the model interprets rules focusing on shape or color.

Pitfalls and Observations. Despite solid overall zero-shot performance, we observed a pronounced drop in accuracy on highly complex or multi-step constraints. This result warns that naive assumption of broad generalization can be problematic, especially when rules require multi-hop reasoning. Furthermore, certain shape-color combinations not encountered in training led to misclassifications, revealing a notable pitfall in real-world scenarios where domain expansions frequently occur.

Hyperparameters. Our neural backbone is a 2-layer MLP with hidden dimension 128, ReLU activations, and dropout of 0.1. We train for 20 epochs, using Adam with learning rate 1×10^{-3}. During inference, the symbolic engine dynamically binds rule variables to embeddings, incurring minimal overhead.

5 CONCLUSION

We introduced a neural-symbolic approach for zero-shot Synthetic PolyRule Reasoning and analyzed real-world pitfalls that emerge when deploying such a system on novel constraints. Our findings confirm that explicit rule representations can enable flexible zero-shot adaptation, but complexity, coverage gaps, and domain expansions pose substantial challenges. Future investigations could explore larger rule spaces, incorporate more adaptive symbolic transformations, and consider robust ways to preempt or mitigate performance drops on unforeseen rules.

REFERENCES

SUPPLEMENTARY MATERIAL

In this appendix, we describe additional experimental details and supplementary results not covered in the main text. Where possible, we provide deeper insights into the pitfalls our model encountered.

A ADDITIONAL IMPLEMENTATION DETAILS

We implement the model in PyTorch. For the MLP backbone, each layer consists of 128 hidden units, ReLU activation, and a dropout probability of 0.1. We found that maintaining a small hidden dimension helps the symbolic engine remain robust to overfitting on certain rule templates.

B FURTHER ABLATIONS

We tested the effect of reducing the number of known symbolic templates used in training. Results showed that decreasing training coverage significantly increases errors on unseen rules, emphasizing the importance of carefully balancing the symbolic complexity and coverage in practice.