Project Proposal

Testing DeepProbLog on Sudoku

by Rohan Sambidi

Abstract

DeepProbLog is a Neuro-Symbolic (Ne-Sy) paradigm that incorporates deep learning into an existing probabilistic logic programming language, ProbLog, through neural predicates. DeepProbLog has been tested on various complex problems that involve learning and reasoning such as MNIST addition, sorting, word algebra problems, etc. In solving these problems, DeepProbLog achieved state-of-the-art results. However, to the best of my knowledge, there has been no implementation of DeepProbLog for solving Sudoku. So, for this project, I will be testing DeepProbLog on Sudoku. I aim to analyze the performance of learning and inference by DeepProbLog for two tasks: classifying visual Sudoku, and solving visual Sudoku. My goal is to gain deeper insights into the integration of neural nets with ProbLog, and compare DeepProbLog with various Ne-Sy frameworks. This would require implementing a Convolutional Neural Network for digit recognition and integrating it with a Probabilistic Graphical Model (like Bayesian Network), for solving, using DeepProbLog.

Introduction

Neural networks are best known for their incredible capability of low-level perception, whereas symbolic AI is capable of high-level reasoning. With incredible strides being made in the field of deep learning, there are still numerous drawbacks that have still not been addressed. Some of these are based on explainability of the deep learning black-box, excessive computation usage, and requirement of high amounts of data for diminishing performance improvements. Symbolic AI (logical and probabilistic representations) can counter these issues but is restricted by the necessity of manually building knowledge graphs. Ne-Sy AI integrates these opposing ends of the spectrum and tries to be the best of both worlds by offering performance comparable to deep learning models, while also being explainable. Integration of neural networks and symbolic reasoning for machine level perception is being studied for over a decade now, and has proven to give promising results.

Various studies have proven that incorporating explainability into deep learning can yield problem-specific solutions, but is usually challenging and would lack scalability and generalizability. Hence, Manhaeve at al. approached the problem from the other end *[Manhaeve et al., 2018]*. They tried to incorporate perceptibility into symbolic AI. This resulted in DeepProbLog, a Ne-Sy paradigm capable of representing and reasoning.

Training AI to solve puzzles is an age-old approach for validation. Puzzles like Sudoku have proven to be challenging for deep learning models, since they lack explainability. Since Ne-Sy

frameworks can incorporate reasoning, testing these models on Sudoku would be a good approach to validate them. However, DeepProbLog was not tested on Sudoku in the original work by Manhaeve at al. Manhaeve et al. released an extended version of the original DeepProbLog publication [Manhaeve et al., 2021] that doesn't included tests on Sudoku either. On the other hand, Sudoku has been used for testing by the authors of other Ne-Sy paradigms. Few examples include NeuPSL, NeurASP, and SARNet. NeuPSL *[Pryor et al., 2022]* was tested on visual Sudoku classification and achieved an accuracy of 85.05 ± 02.65% when trained on 200 puzzles. NeurASP *[Yang et al., 2020]* was tested on solving Sudoku and achieved an accuracy of 100% on when trained on 25 puzzles. SATNet *[Wang et al., 2019]* was also tested on solving Sudoku and achieved an accuracy of 63.2% on visual Sudoku and 98.3% on original Sudoku, when trained on 9000 puzzles. Since many paradigms use Sudoku for validation, testing DeepProbLog on Sudoku would allow fair comparison with those paradigms.

To the best of my knowledge, there have been no studies on testing DeepProbLog on Sudoku. Hence I chose to test DeepProbLog on Sudoku for this project. I propose two problems for testing that involves perception and reasoning:

T1- Classifying visual Sudoku: Given a Sudoku board, determine whether the puzzle is complete (solved) or incomplete (unsolved).

T2- Solving visual Sudoku: Given an incomplete Sudoku grid, solve the puzzle.

Since Ne-Sy AI is a fairly new field, there is no proper benchmark to evaluate Ne-Sy frameworks *[Augustine et al., 2022]*. Augustine et al. created a database (ViSudo-PC) for the test on classifying visual Sudoku which can serve a benchmark for this particular test. I will use this dataset to evaluate DeepProbLog for task T1. For task T2, I'll use the MNIST SudokuSolver dataset (which was used by Wang et al. to test SATNet).

Group Members

Rohan Sambidi

I have worked on a course project on Ne-Sy AI before. The project was aimed at comparing performance of Ne-Sy models with deep learning models over varied data sizes. The idea was that, since Ne-Sy paradigms utilize the reasoning capabilities of symbolic AI, unlike traditional deep learning models, training on smaller datasets should be sufficient to achieve satisfactory results.

Measure of Success

Since the whole project is focused on modeling a Sudoku solver in DeepProgLog, the baseline success would be to have a model that can at least percept the Sudoku board and try solving it. Representing visual Sudoku symbolically is necessary to proceed forward.

- Baseline: The model is able to recognize the digits and represent them in ProbLog.
- Stretch: The model is able to recognize the digits *correctly*, even with noisy input, understand whether the puzzle is solved, and, if it's not solved, then solve it. A further stretch

would be to utilize the capabilities of the probabilistic reasoner to *correctly* solve a puzzles of various sizes, even when originally trained on the classic 9x9 Sudoku grid.

Preliminary Plan

I will follow the timeline below to complete this project in a timely manner:

- First week of March
  - Design a neural network for digit classification
  - Get familiar with DeepProbLog environment
- Last week of March (Check-in on March 31)
  - First test run of DeepProbLog on smaller experiments like MNIST addition
  - Test DeepProbLog for test T1
  - Reflect on the feasibility of modeling Sudoku in ProbLog
- Third week of April (Presentation on April 19)
  - Evaluate DeepProbLog based on the test results
  - Compare the performance with other paradigms
  - Complete the project and present the findings
- Fourth week of April (Report Submission on April 28)

References

I. R. Manhaeve, S. Dumancic, A. Kimmig, T. Demeester, & L. D. Raedt. DeepProbLog: Neural probabilistic logic programming. In Proceedings of Advances in Neural Information Processing Systems, p. 3749–3759, 2018. https://doi.org/10.48550/arXiv.1805.10872

II. R. Manhaeve, S. Dumancic, A. Kimmig, T. Demeester, & L. D. Raedt. Neural probabilistic logic programming in DeepProbLog . In Artificial Intelligence, vol. 298, p. 103504, Elsevier BV, 2021. https://doi.org/10.48550/arXiv.1907.08194

III. E. Augustine, C. Pryor, C. Dickens, J. Pujara, W. Y. Wang, & L. Getoor. Visual Sudoku Puzzle Classification: A Suite of Collective Neuro-Symbolic Tasks. In NeSy 2022, 16th International Workshop on Neural-Symbolic Learning and Reasoning, 2022. https://ceur-ws.org/Vol-3212/paper2.pdf

IV. Pryor, C., Dickens, C., Augustine, E., Albalak, A., Wang, W., & Getoor, L, NeuPSL: Neural Probabilistic Soft Logic, arXiv, 2022. https://doi.org/10.48550/arXiv.2205.14268

V. Z. Yang, A. Ishay, & J. Lee, NeurASP: Embracing Neural Networks into Answer Set Programming, International Joint Conferences on Artificial Intelligence Organization, 2020. http://dx.doi.org/10.24963/ijcai.2020/243

VI. P. Wang, P. L. Donti, B. Wilder, & Z. Kolter, SATNet: Bridging Deep Learning and Logical Reasoning Using a Differentiable Satisfiability Solver, International Conference on Machine Learning, 2019. https://doi.org/10.48550/arXiv.1905.12149