Independence is not the issue

Exploring common assumptions in Neurosymbolic AI

Håkan Karlsson Faronius [1] Pedro Zuidberg dos Martires [1] Luc De Raedt [1,2]

[1]Örebro University [2]KU Leuven

Abstract

In this work we wish to refute some of the claims made in [2], specifically that the conditional independence assumption in Neurosymbolic AI causes models to exhibit a deterministic bias. We will here show that it is in fact the determinism that is the cause of this deterministic bias.

Traffic Light Example

We will here introduce neurosymbolic AI by using a traffic light example from [2].

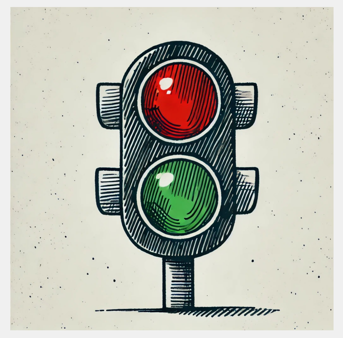

Figure 1. A traffic light.

This traffic light has a red light and a green light. The traffic light can be in one of the following states:

- (red, ¬ green)
- (¬ red, green)
- (¬ red, ¬ green)

The **logical constraint** φ is satisfied if it is one of these "**possible worlds**".

Neurosymbolic AI

In Neurosymbolic AI we use neural networks to predict probability distributions for the data. In this case we would have the following networks:

- **Network 1** predicts if the **red light** is **on/off**.
- **Network 2** predicts if the **green light** is **on/off**.

The neural networks are called **propositional neural predicates**, and are denoted as w_1 and w_2 respectively. Together they form possible worlds $\mathbf{w} = (w_1, w_2)$ if they satisfy the logical constraint.

The probability that they satisfy a given logical constraint φ can be calculated as

$$\mathbb{P}_\theta(\varphi = 1|\mathbf{x}) = \sum_{\mathbf{w}\in\{0,1\}^n} \mathbb{P}_\theta(\varphi|\mathbf{w})\mathbb{P}_\theta(\mathbf{w}|\mathbf{x}) \quad (1)$$

where 1 and 0 denotes true and false respectively and the input data is denoted as $\mathbf{x}$.

The conditional independence assumption in Neurosymbolic AI

In Neurosymbolic AI it is common to make the following **conditional independence assumption** that the propositional neural predicates w_i, w_j are conditionally independent given input $\mathbf{x}$, for all $i \neq j$, mathematically

$$\mathbb{P}_\theta(\mathbf{w}|\mathbf{x}) = \prod_{i=1}^{n} \mathbb{P}_\theta(w_i|\mathbf{x}) \quad (2)$$

Intuitively, *predicting each propositional neural predicate does not affect the predictions of any others.*

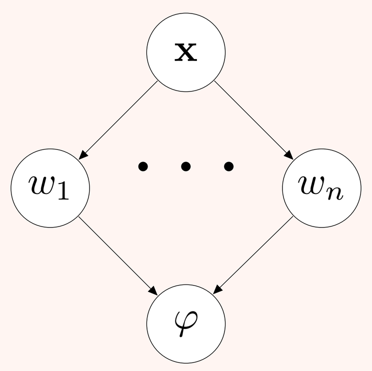

Figure 2. A Bayesian network visualization.

The Article by Emile van Krieken et al[2]

[2] make several claims about the consequences of using the conditional independence assumption. Their main point is that independence, together with determinism (specifically determinism in the sense that $\mathbb{P}_{\theta_1}(\varphi|\mathbf{w}) = 1$)) causes neurosymbolic models to become overly confident in their predictions, which is referred to as a **deterministic bias**.

$$\text{Independence} \wedge \text{Determinism} \implies \text{Deterministic Bias}$$

Data for traffic lights experiments

For the experiments we construct "traffic lights" by using 1's and 0's from the MNIST dataset, representing on and off respectively for each light, , , , representing each of the possible worlds.

Experiments in DeepProbLog

We can verify these findings by using DeepProbLog[1] and modeling the traffic light from before.

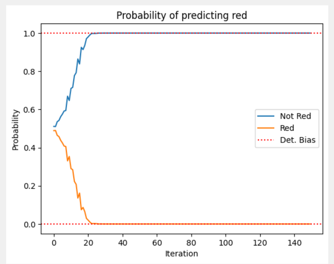

Figure 3. Probability of the red light being on/off during training.

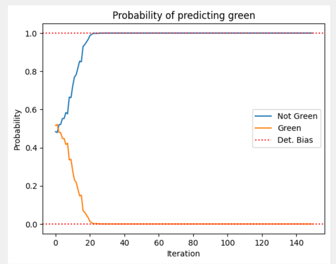

Figure 4. Probability of the green light being on/off during training.

This leads us to conclude that **independence** and **determinism** does seem to cause a **deterministic bias**.

Assuming Conditional Dependence

Instead of assuming independence as before, we can now instead look at models where the **propositional neural predicates** are **conditionally dependent**. For the traffic light example this can be implemented by making the prediction of green w_2 conditionally dependent on the prediction of the red light w_1 given input $\mathbf{x}$.

$$\mathbb{P}_\theta(\mathbf{w}|\mathbf{x}) = \mathbb{P}_\theta(w_1|\mathbf{x})\mathbb{P}_\theta(w_2|w_1, \mathbf{x}) \quad (3)$$

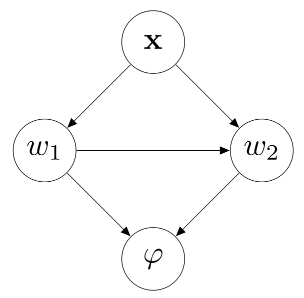

Figure 5. A Bayesian network visualization.

Experiments in DeepProbLog

We can implement the conditional dependence in DeepProbLog[1] by conditioning on the prediction of the red light being on or off and then training a neural network to predict if the green light is on/off for each outcome of the first prediction.

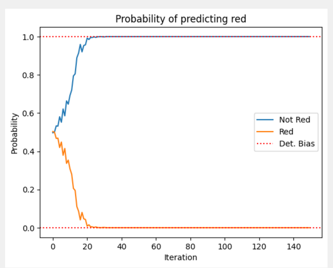

Figure 6. Probability of the red light being on.

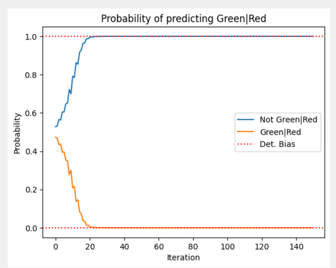

Figure 7. Probability of green light being on, given that the red light is on.

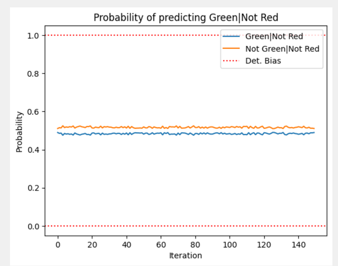

Figure 8. Probability of green light being on, given that the red light is off. Always true, so it is never trained.

From these experiments however it should be apparent that **dependence** and **determinism** also seem to cause a **deterministic bias**.

This would lead us to conclude

$$\text{Dependence} \wedge \text{Determinism} \implies \text{Deterministic Bias}$$

Nondeterminism

We can also consider the case when we assume independence and non-determinism and conclude experimentally that it does not result in a deterministic bias.

Experiments in DeepProbLog

We can test this by using DeepProbLog[1] and modeling the traffic light from before. Here the logical condition is intentionally false in 60% of cases.

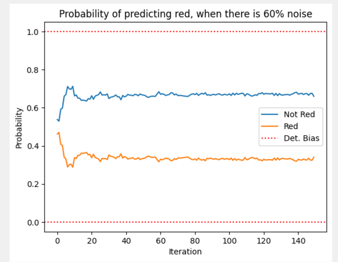

Figure 9. Probability of the red light being on/off during training.

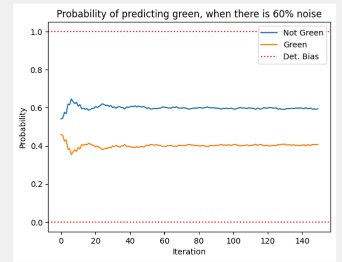

Figure 10. Probability of the red light being on/off during training.

This leads us to conclude that **independence** and **non-determinism** does **not** lead to a **deterministic bias**.

$$\text{Independence} \wedge \neg\text{Determinism} \not\Rightarrow \text{Deterministic Bias}$$

Conclusion

Based on these experimental findings, the fact that a probability distribution is either independent or dependent, we can **conclude** that it is in fact

$$\text{Determinism} \implies \text{Deterministic Bias}$$

and the conditional independence assumption in neurosymbolic AI has nothing to do with it.

References

[1] Robin Manhaeve, Sebastijan Dumančić, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. Neural probabilistic logic programming in deepproblog. *Artificial Intelligence*, 298:103504, 2021. ISSN 0004-3702. doi: https://doi.org/10.1016/j.artint.2021.103504. URL https://www.sciencedirect.com/science/article/pii/S0004370221000552.

[2] Emile van Krieken, Pasquale Minervini, Edoardo M. Ponti, and Antonio Vergari. On the independence assumption in neurosymbolic learning. In *ICML*. OpenReview.net, 2024. URL http://dblp.uni-trier.de/db/conf/icml/icml2024.html#KriekenMPV24.

hakan.karlsson-faronius@oru.se