

DeepMind

Model-Value Inconsistency

Angelos Filos*, Eszter Vértés*, **Zita Marinho***, Gregory Farquhar, Diana Borsa,
Abram Friesen, Feryal Behbahani, Tom Schaul, André Barreto, Simon Osindero

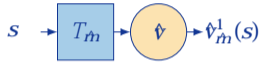
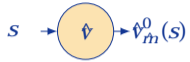
ICML 2022



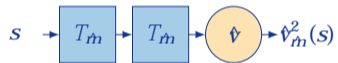
Key Observation



Key Observation



Key Observation



Key Observation

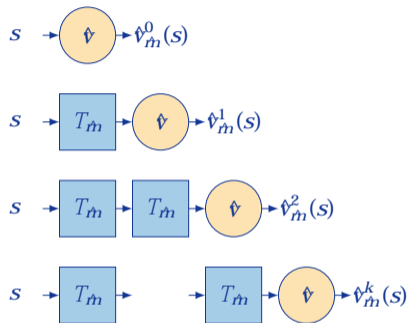


Figure 1: Ensemble value predictions from a single model \hat{m} and value function \hat{v} .

Key Observation

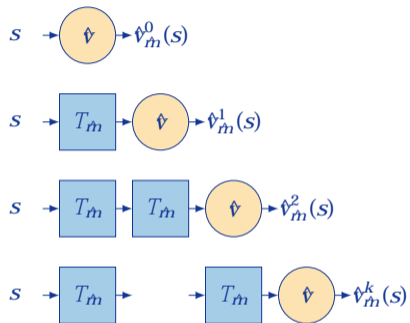


Figure 1: Ensemble value predictions from a single model m and value function \hat{v} .

Implicit Value Ensemble (IVE)

$$f_{\hat{v}_m^i}^k, \quad \frac{f_{\hat{v}}; T_m \hat{v}; \dots; (T_m)^k \hat{v} g}{k+1 \text{ value predictions}}$$

Key Observation

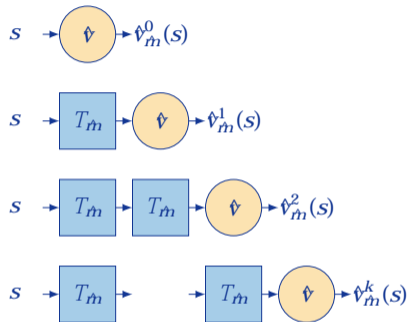


Figure 1: Ensemble value predictions from a single model m and value function \hat{v} .

Implicit Value Ensemble (IVE)

$$f_{\hat{v}_m^i g_{i=0}^k}, \quad \frac{f_{\hat{v}; T_m \hat{v}; \dots; (T_m)^k \hat{v}} g}{k+1 \text{ value predictions}}$$

Model-Value Inconsistency

Implicit Value Ensemble disagreement, e.g.,

$$-\text{IVE}(k), \quad \exists \hat{v} \neq f_{\hat{v}_m^i g_{i=0}^k}$$

ì È ¥â

Research Goal

Determine how effective model-value inconsistency is as a signal for epistemic uncertainty.

ì È ¥â

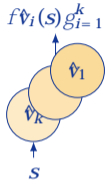
Research Goal

Determine how effective model-value inconsistency is as a signal for epistemic uncertainty.

Key Takeaway

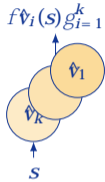
, ÔâÎ Ô¥È ocEY Qí i »æcYâcÈ EBâÔâ »í ¶ cĩ æÍ ocÈ ÔûcÈ ©µÍ Ýi»Œ æ

Value Uncertainty Quantification

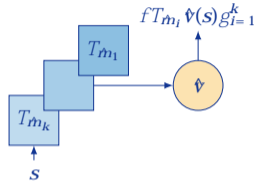


(a) Value Ensemble (EVE)

Value Uncertainty Quantification

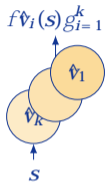


(a) Value Ensemble (EVE)

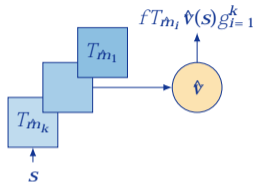


(b) Model Ensemble (EMVE)

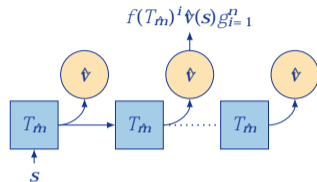
Value Uncertainty Quantification



(a) Value Ensemble (EVE)



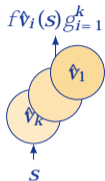
(b) Model Ensemble (EMVE)



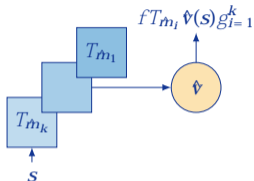
(c) Implicit Value Ensemble (IVE)

Figure 2: Value uncertainty quantification in scalable epistemic uncertainty-aware RL agents.

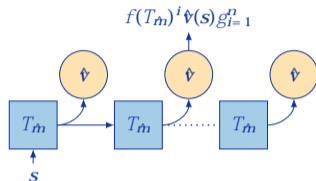
Value Uncertainty Quantification



(a) Value Ensemble (EVE)



(b) Model Ensemble (EMVE)



(c) Implicit Value Ensemble (IVE)

Figure 2: Value uncertainty quantification in scalable epistemic uncertainty-aware RL agents.

Key Difference

IVE relies on an ensemble of a learned model \hat{m} and a value function \hat{v} .

IVE as proxy for Epistemic Uncertainty

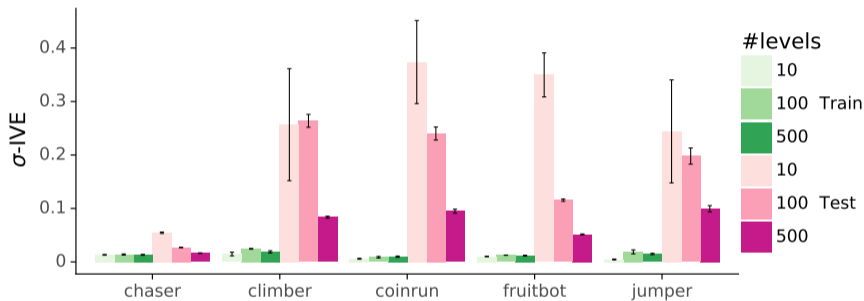


Figure 3: σ -IVE(5) computed using the model of the Muesli agent while evaluating on both training (green) and unseen test levels (pink). Bars, error-bars show mean and ste across 3 seeds.



Conclusion

Investigate the qualitative and quantitative properties of IVE:

- Disagreement among IVE members is an effective signal for epistemic uncertainty
- it is a good signal to guide exploration
- IVEs can robustify value-based planning, and help deal with distribution shifts.

