Learning Tree Interpretation from Object Representation for Deep Reinforcement Learning

Guiliang Liu (Presenter),

Xiangyu Sun, Oliver Schulte, and Pascal Poupart

Presented at NEURAL INFORMATION PROCESSING SYSTEMS

VECTOR INSTITUTE | INSTITUT VECTEUR

UNIVERSITY OF WATERLOO

SFU SIMON FRASER UNIVERSITY

Introduction

Problem Definition:

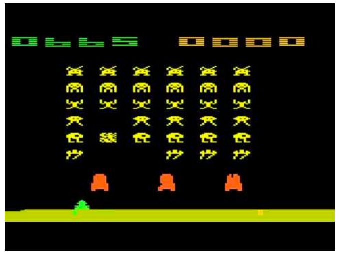
Target: Learning Tree Interpretation for DRL Agents in complex environments.

Problem Definition:

Target: Learning Tree Interpretation for DRL Agents in complex environments.

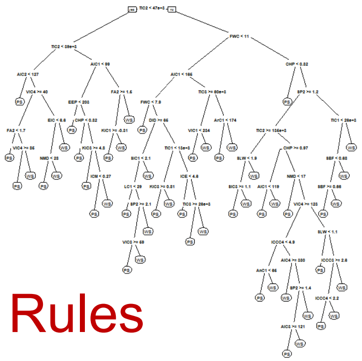
Difficulties & Challenges:

- High-dimensional input space.



$$H = 49{,}152$$
$$(128 * 128 * 3)$$

- Complex Tree Model



Complex Decision Rules

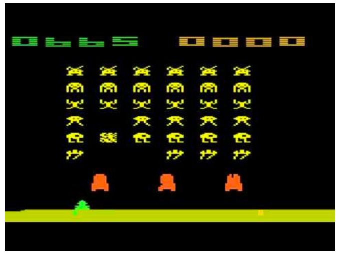
Introduction

Problem Definition:

Target: Learning Tree Interpretation for DRL Agents in complex environments.

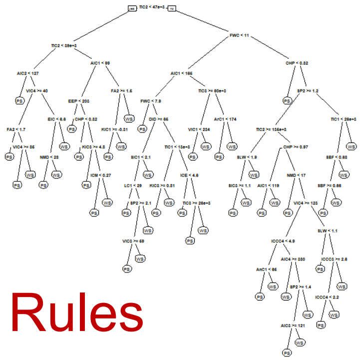
Difficulties & Challenges:

- High-dimensional input space.



$$H = 49{,}152$$
$$(128 * 128 * 3)$$

- Complex Tree Model



Complex Decision Rules

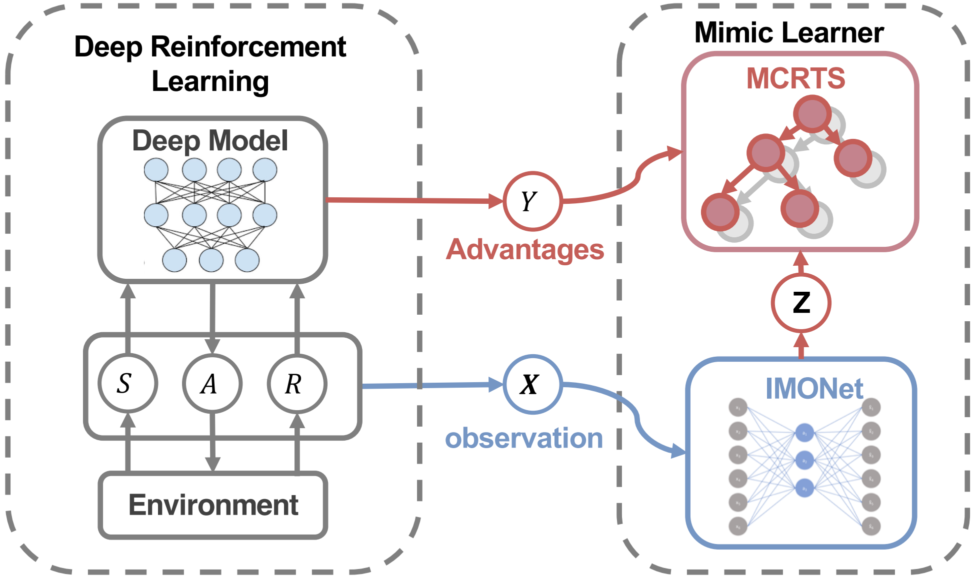
Represent And Mimic Framework (RAMi):

IMONet: Interpretable representation model.

- Learning a disentangled representation.

MCRTS: Interpretable decision model.

- Learning a mimic tree.



Represent And Mimic Framework:

The **Information Bottleneck (IB)** objectives for Representing and Mimicking:

- The IB objective: $\max_{\boldsymbol{\omega}} \left[I_{\boldsymbol{\omega}}(\boxed{\Phi}, \boxed{Y}) - \lambda I_{\boldsymbol{\omega}}(\Phi, \boxed{X}) \right]$

Mimic Tree Mimic Targets Input Signals

Represent And Mimic Framework:

The **Information Bottleneck (IB)** objectives for Representing and Mimicking:

- The IB objective: $\max_{\boldsymbol{\omega}} \left[I_{\boldsymbol{\omega}}(\boxed{\Phi}, \boxed{Y}) - \lambda I_{\boldsymbol{\omega}}(\Phi, \boxed{X}) \right]$

  Mimic Tree Mimic Targets Input Signals

- Learn a latent object representation for X.

  $$q(\boldsymbol{\phi}|\boldsymbol{x}_n) = \int q(\boldsymbol{\phi}|\boldsymbol{z})q(\boldsymbol{z}|\boldsymbol{x}_n)\mathrm{d}\boldsymbol{z}$$

Represent And Mimic Framework:

The **Information Bottleneck (IB)** objectives for Representing and Mimicking:

- The IB objective: $\max_{\boldsymbol{\omega}} \left[I_{\boldsymbol{\omega}}(\underbrace{\boxed{\Phi}}_{\text{Mimic Tree}}, \underbrace{\boxed{Y}}_{\text{Mimic Targets}}) - \lambda I_{\boldsymbol{\omega}}(\Phi, \underbrace{\boxed{X}}_{\text{Input Signals}}) \right]$

- Learn a latent object representation for X.

$$q(\boldsymbol{\phi}|\boldsymbol{x}_n) = \int q(\boldsymbol{\phi}|\boldsymbol{z})q(\boldsymbol{z}|\boldsymbol{x}_n)\mathrm{d}\boldsymbol{z}$$

- Computing the IB objective is impractical, so
  - Approximate $p(x, y)$ with the empirical distribution $1/N \sum_{n=1}^{N} \delta_{y_n}(y)\delta_{\boldsymbol{x}_n}(\boldsymbol{x})$.
  - Apply the deep variational method for IB [30].

[30] Alexander A. Alemi, Ian Fischer, Joshua V. Dillon, and Kevin Murphy. Deep variational information bottleneck. In 5th International Conference on Learning Representations, ICLR 2017.

Represent And Mimic Framework:

The **Information Bottleneck (IB)** objectives for Representing and Mimicking:

- The IB objective: $\max_{\boldsymbol{\omega}} \left[I_{\boldsymbol{\omega}}(\boxed{\Phi}, \boxed{Y}) - \lambda I_{\boldsymbol{\omega}}(\Phi, \boxed{X}) \right]$

  Mimic Tree Mimic Targets Input Signals

- Learn a latent object representation for X.

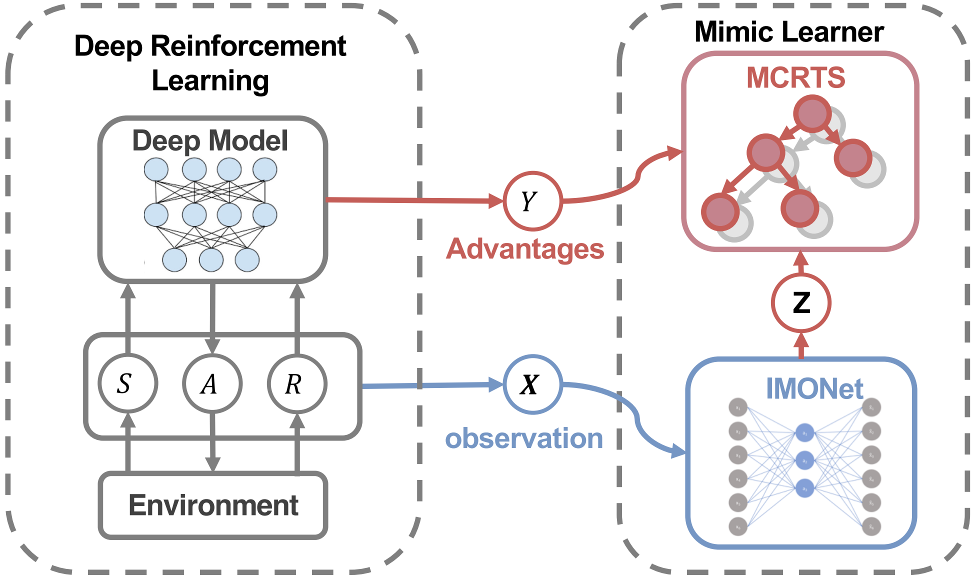
  $$q(\phi|\boldsymbol{x}_n) = \int q(\phi|\boldsymbol{z})q(\boldsymbol{z}|\boldsymbol{x}_n)\mathrm{d}\boldsymbol{z}$$

- Computing the IB objective is impractical, so
  - Approximate $p(x, y)$ with the empirical distribution $1/N \sum_{n=1}^{N} \delta_{y_n}(y)\delta_{\boldsymbol{x}_n}(\boldsymbol{x})$.
  - Apply the deep variational method for IB [30].

- The lower bound of IB objective is:

$$\frac{1}{N}\sum_{n=1}^{N}\Big\{\mathbb{E}_{q(\boldsymbol{z}|\boldsymbol{x}_n)}[\log p(\boldsymbol{x}_n|\boldsymbol{z})] - \lambda \mathcal{D}_{KL}[q(\boldsymbol{z}|\boldsymbol{x}_n)\|p_0(\boldsymbol{z})] - \mathbb{E}_{q(\boldsymbol{z}|\boldsymbol{x}_n)}\Big[\mathbb{E}_{q(\phi|\boldsymbol{z})}\Big(\mathcal{L}_q(y_n) + \lambda\mathcal{L}_p(\phi)\Big) - \lambda H[q(\phi|\boldsymbol{z})]\Big]\Big\}$$

$$= \textit{ELBo objective} + \textit{IB-MDL objective} + \textit{Entropy Regularizer}$$



[30] Alexander A. Alemi, Ian Fischer, Joshua V. Dillon, and Kevin Murphy. Deep variational information bottleneck. In 5th International Conference on Learning Representations, ICLR 2017.

Learning Object Representation

Identifiable Multi-Objects Network (IMONet):

Motivation 1: Learning a disentangled representation.

$p(Z) = \prod_{d=1}^{D} p(Z_d)$ (Following the VAE framework.)

Learning Object Representation

Identifiable Multi-Objects Network (IMONet):

Motivation 1: Learning a disentangled representation.

$p(Z) = \prod_{d=1}^{D} p(Z_d)$ (Following the VAE framework.)

Motivation 2: Learning an identifiable representation.

- Unidentifiable → different factorizations for the same inputs.
- Identifiable VAE (IVAE) [33] → conditionally factored prior.

$$p(Z|A,R) = \prod_d p(Z_d|A,R)$$

- $Z_{d,k}$ captures an independent factor of object variations

[33] Ilyes Khemakhem, Diederik P. Kingma, Ricardo Pio Monti, and Aapo Hyvärinen. Variational autoencoders and nonlinear ICA: A unifying framework. AISTATS 2020.

Learning Object Representation

Identifiable Multi-Objects Network (IMONet):

Motivation 1: Learning a disentangled representation.

$$p(Z) = \prod_{d=1}^{D} p(Z_d) \text{ (Following the VAE framework.)}$$

Motivation 2: Learning an identifiable representation.

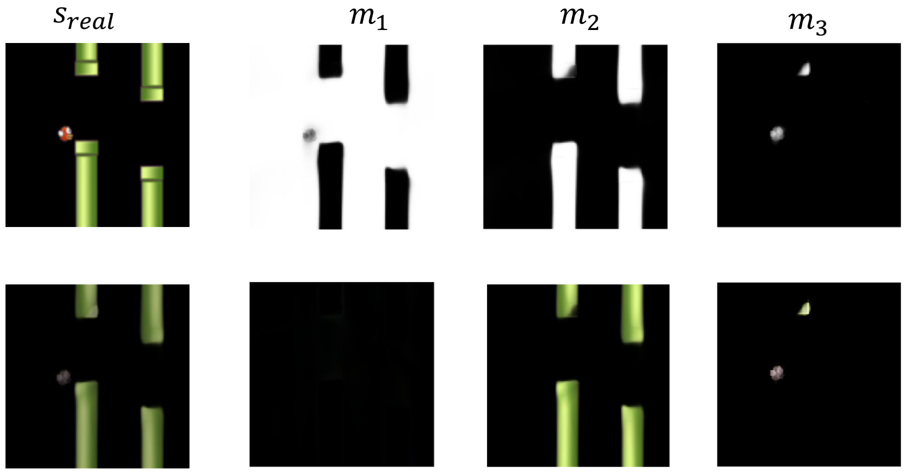
- Unidentifiable → different factorizations for the same inputs.
- Identifiable VAE (IVAE) [33] → conditionally factored prior.

$$p(Z|A,R) = \prod_d p(Z_d|A,R)$$

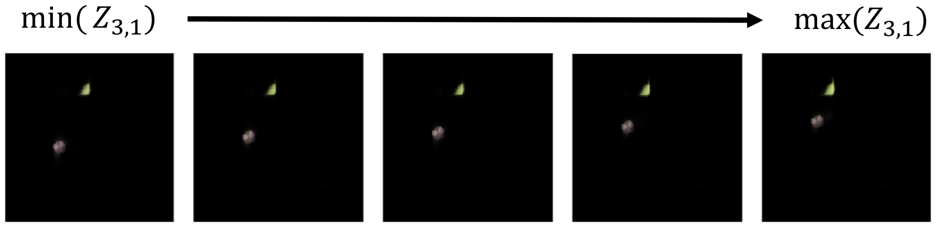
Motivation 3: Learning an interpretable representation.

- IMonet (follows Monet [34])learns a symbolic abstraction of state space by representing object variations.
- $Z_{d,k}$ captures an independent factor of object variations

a) Decomposes a state into objects.



b) Represent an object variation with a latent variable.



[33] Ilyes Khemakhem, Diederik P. Kingma, Ricardo Pio Monti, and Aapo Hyvärinen. Variational autoencoders and nonlinear ICA: A unifying framework. AISTATS 2020.

[34] Christopher P. Burgess, et,al. Monet: Unsupervised scene decomposition and representation. CoRR, abs/1901.11390, 2019.

Learning Mimic Tree Interpretations

Inferring Mimic Trees with IB-MDL

$$\mathbb{E}_{q(\boldsymbol{\phi}|\boldsymbol{z})}\left(\mathcal{L}_q(y_n) + \lambda\mathcal{L}_p(\boldsymbol{\phi})\right) - \lambda H[q(\boldsymbol{\phi}|\boldsymbol{z})]$$

Learning Mimic Tree Interpretations

Inferring Mimic Trees with IB-MDL

$$\mathbb{E}_{q(\boldsymbol{\phi}|\boldsymbol{z})}\left(\mathcal{L}_q(y_n) + \lambda\mathcal{L}_p(\boldsymbol{\phi})\right) - \lambda H[q(\boldsymbol{\phi}|\boldsymbol{z})]$$

Minimize $\mathcal{L}_p(\boldsymbol{\phi})$: the description length of encoding the tree structure.

- Convert the binary tree structure to a string [42].

Proposition 1 *Given a regression tree with L splits, the total cost (in bits) of describing the tree structure with the string encoding method is:*

$$\mathcal{L}_p(\phi) = \log\frac{(2L-1)^2}{L^{\frac{3}{2}}(L-1)^{\frac{1}{2}}} + (2L-1)H(\frac{L}{2L-1}) + O(L^{-1}) \qquad (5)$$

[42] J. Ross Quinlan and Ronald L. Rivest. Inferring decision trees using the minimum description length principle. Information and Computation, 80(3):227–248, 1989.

Learning Mimic Tree Interpretations

Inferring Mimic Trees with IB-MDL

$$\mathbb{E}_{q(\boldsymbol{\phi}|\boldsymbol{z})}\Big(\mathcal{L}_q(y_n)+\lambda\mathcal{L}_p(\boldsymbol{\phi})\Big)-\lambda H[q(\boldsymbol{\phi}|\boldsymbol{z})]$$

Minimize $\boldsymbol{\mathcal{L}_p(\phi)}$: the description length of encoding the tree structure.

- Convert the binary tree structure to a string [42].

Proposition 1 *Given a regression tree with L splits, the total cost (in bits) of describing the tree structure with the string encoding method is:*

$$\mathcal{L}_p(\phi)=\log\frac{(2L-1)^2}{L^{\frac{3}{2}}(L-1)^{\frac{1}{2}}}+(2L-1)H(\frac{L}{2L-1})+O(L^{-1}) \quad (5)$$

Minimize $\boldsymbol{\mathcal{L}_q(y_n)}$: the description length of encoding the exceptions.

- Model the leaf prediction with a Gaussian distribution.
- Apply the log-variance as the cost of encoding exceptions

[42] J. Ross Quinlan and Ronald L. Rivest. Inferring decision trees using the minimum description length principle. Information and Computation, 80(3):227–248, 1989.

Learning Mimic Tree Interpretations

Inferring Mimic Trees with IB-MDL

$$\mathbb{E}_{q(\phi|\boldsymbol{z})}\Big(\mathcal{L}_q(y_n) + \lambda\mathcal{L}_p(\phi)\Big) - \lambda H[q(\phi|\boldsymbol{z})]$$

Minimize $\mathcal{L}_p(\phi)$: the description length of encoding the tree structure.

- Convert the binary tree structure to a string [42].

Proposition 1 *Given a regression tree with L splits, the total cost (in bits) of describing the tree structure with the string encoding method is:*

$$\mathcal{L}_p(\phi) = \log \frac{(2L-1)^2}{L^{\frac{3}{2}}(L-1)^{\frac{1}{2}}} + (2L-1)H\left(\frac{L}{2L-1}\right) + O(L^{-1}) \tag{5}$$

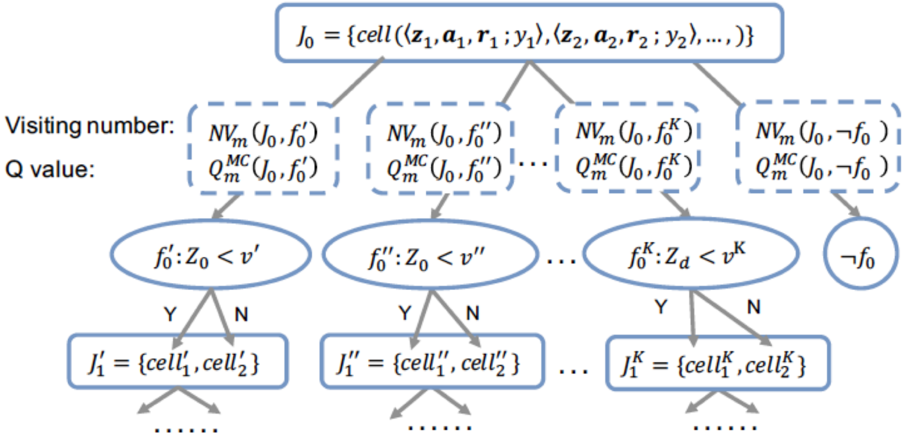
Minimize $\mathcal{L}_q(y_n)$: the description length of encoding the exceptions.

- Model the leaf prediction with a Gaussian distribution.
- Apply the log-variance as the cost of encoding exceptions

Monte-Carlo Regression Tree Search (MCRTS)

- Learns a distribution of mimic trees based on the latent features from the object representation.
- The reward is defined by:

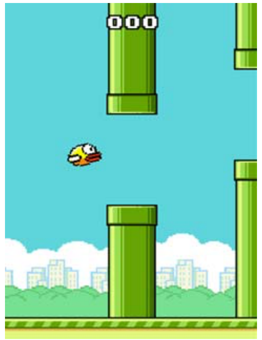
$$r^{MC}(J_{leaf}) = -\mathcal{L}_q(y_n) - \lambda\mathcal{L}_p(\phi)$$



[42] J. Ross Quinlan and Ronald L. Rivest. Inferring decision trees using the minimum description length principle. Information and Computation, 80(3):227–248, 1989.

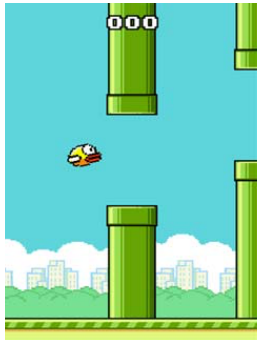
Empirical Evaluation

Environments: 1) Flappy Bird 2) Space Invaders 3) Assault



Empirical Evaluation

Environments: 1) Flappy Bird 2) Space Invaders 3) Assault



Baselines:

- Tree learners:
  a) CART: Classification And Regression Tree.
  b) VIPER: Q-dagger based imitation learner.
  c) M5-RT/MT: learn the regression tree or the model tree based on M5 algorithm
  d) GM/VR-LMT: Linear Model Tree based on Variance Reduction (VR) and Gaussian Mixture (GM) for feature selection.
- Representation Learners:
  a) Classic VAE

Fidelity versus Simplicity

Metrics:

- **Fidelity**: Variance Reduction (VR)
- **Simplicity**: Leaf Numbers.
- **Fidelity v.s., Simplicity:** VR Per-Leaf (VR-PL)

	Flappy Bird			Space Invaders			Assault		
Method	VR	VR-PL	Leaf	VR	VR-PL	Leaf	VR	VR-PL	Leaf
Cart	8.51E-2	8.43E-5	1007	4.96E-2	7.02E-5	705	4.79E-2	7.46E-5	642
VIPER	8.57E-2	1.88E-4	453	4.63E-2	8.80E-5	525	5.28E-2	8.09E-5	653
M5-RT	**9.59E-2**	8.37E-5	1144	4.54E-2	2.92E-5	1558	4.37E-2	2.73E-5	1605
M5-MT	9.56E-2	1.55E-4	612^{w+}	1.60E-2	1.23E-5	1303^{w+}	3.42E-2	2.54E-5	1351^{w+}
GM-LMT	8.99E-2	2.99E-4	303^{w+}	2.07E-2	8.32E-5	249^{w+}	5.55E-2	1.83E-4	307^{w+}
VR-LMT	8.46E-2	5.36E-4	157^{w+}	2.65E-2	1.61E-4	166^{w+}	5.80E-2	1.98E-4	291^{w+}
VAE+CART	7.25E-2	3.44E-4	212	3.99E-2	7.86E-5	507	5.15E-2	1.16E-4	448
VAE+VIPER	7.63E-2	5.32E-4	143	4.12E-2	9.89E-5	417	4.57E-2	1.29E-4	356
VAE+GM-LMT	6.35E-2	3.51E-4	180^{w+}	3.39E-2	2.75E-4	123^{w+}	4.20E-2	1.44E-5	293^{w+}
VAE+VR-LMT	7.95E-2	5.12E-4	154^{w+}	3.52E-2	2.08E-4	171^{w+}	5.10E-2	1.99E-4	258^{w+}
VAE+MCRTS	7.83E-2	1.27E-3	61	4.82E-2	5.66E-4	85	6.58E-2	7.75E-4	85
IMONet+CART	8.23E-2	4.02E-4	204	5.21E-2	1.38E-4	375	5.67E-2	1.81E-4	315
IMONet+VIPER	8.50E-2	4.48E-4	191	5.26E-2	1.69E-4	313	6.05E-2	1.90E-4	319
IMONet+GM-LMT	7.87E-2	3.74E-4	212^{w+}	4.79E-2	3.23E-4	149^{w+}	5.45E-2	2.15E-4	256^{w+}
IMONet+VR-LMT	8.21E-2	7.16E-4	115^{w+}	4.54E-2	3.79E-4	120^{w+}	6.03E-2	2.27E-4	268^{w+}
IMONet+MCRTS	8.53E-2	**1.37E-3**	**62**	**5.37E-2**	**7.08E-4**	**76**	**7.53E-2**	**9.07E-4**	**83**

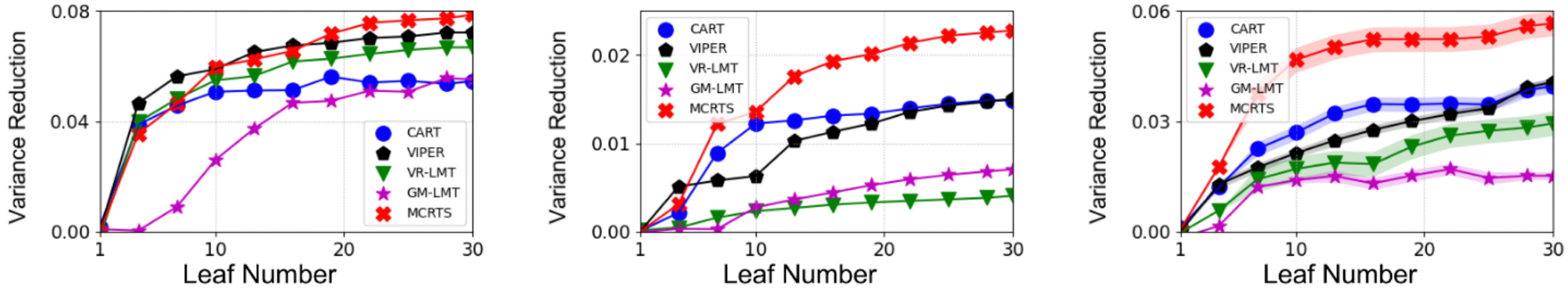
Fidelity versus Simplicity

Metrics:

- **Fidelity**: Variance Reduction (VR)
- **Simplicity**: Leaf Numbers.
- **Fidelity v.s., Simplicity:** VR Per-Leaf (VR-PL)

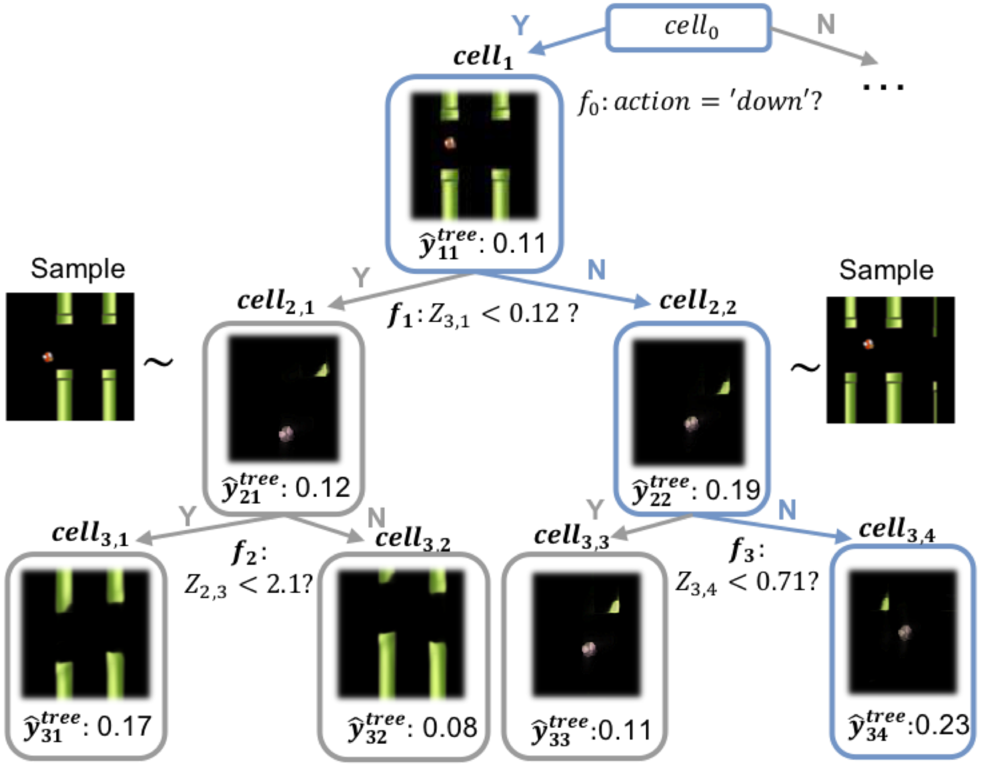
	Flappy Bird			Space Invaders			Assault		
Method	VR	VR-PL	Leaf	VR	VR-PL	Leaf	VR	VR-PL	Leaf
Cart	8.51E-2	8.43E-5	1007	4.96E-2	7.02E-5	705	4.79E-2	7.46E-5	642
VIPER	8.57E-2	1.88E-4	453	4.63E-2	8.80E-5	525	5.28E-2	8.09E-5	653
M5-RT	**9.59E-2**	8.37E-5	1144	4.54E-2	2.92E-5	1558	4.37E-2	2.73E-5	1605
M5-MT	9.56E-2	1.55E-4	612^{w+}	1.60E-2	1.23E-5	1303^{w+}	3.42E-2	2.54E-5	1351^{w+}
GM-LMT	8.99E-2	2.99E-4	303^{w+}	2.07E-2	8.32E-5	249^{w+}	5.55E-2	1.83E-4	307^{w+}
VR-LMT	8.46E-2	5.36E-4	157^{w+}	2.65E-2	1.61E-4	166^{w+}	5.80E-2	1.98E-4	291^{w+}
VAE+CART	7.25E-2	3.44E-4	212	3.99E-2	7.86E-5	507	5.15E-2	1.16E-4	448
VAE+VIPER	7.63E-2	5.32E-4	143	4.12E-2	9.89E-5	417	4.57E-2	1.29E-4	356
VAE+GM-LMT	6.35E-2	3.51E-4	180^{w+}	3.39E-2	2.75E-4	123^{w+}	4.20E-2	1.44E-5	293^{w+}
VAE+VR-LMT	7.95E-2	5.12E-4	154^{w+}	3.52E-2	2.08E-4	171^{w+}	5.10E-2	1.99E-4	258^{w+}
VAE+MCRTS	7.83E-2	1.27E-3	61	4.82E-2	5.66E-4	85	6.58E-2	7.75E-4	85
IMONet+CART	8.23E-2	4.02E-4	204	5.21E-2	1.38E-4	375	5.67E-2	1.81E-4	315
IMONet+VIPER	8.50E-2	4.48E-4	191	5.26E-2	1.69E-4	313	6.05E-2	1.90E-4	319
IMONet+GM-LMT	7.87E-2	3.74E-4	212^{w+}	4.79E-2	3.23E-4	149^{w+}	5.45E-2	2.15E-4	256^{w+}
IMONet+VR-LMT	8.21E-2	7.16E-4	115^{w+}	4.54E-2	3.79E-4	120^{w+}	6.03E-2	2.27E-4	268^{w+}
IMONet+MCRTS	8.53E-2	**1.37E-3**	**62**	**5.37E-2**	**7.08E-4**	**76**	**7.53E-2**	**9.07E-4**	**83**

Leaf-by-Leaf Regression Performance :



Illustrative Examples of Interpretable Mimic Trees

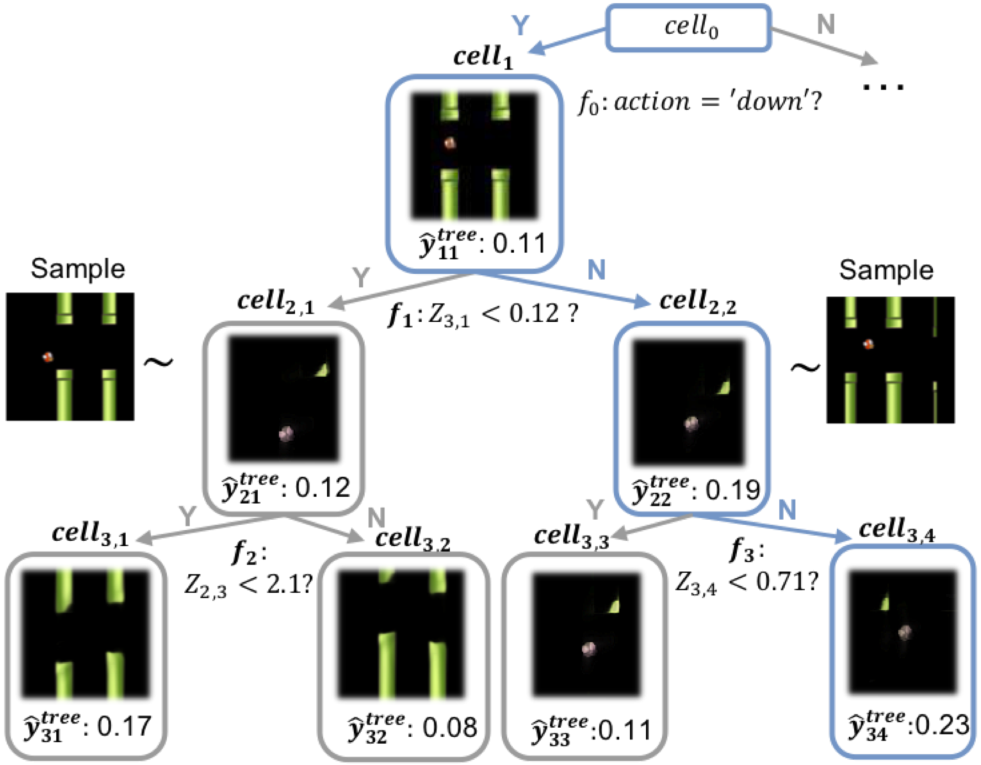
Illustrating the extracted rules :



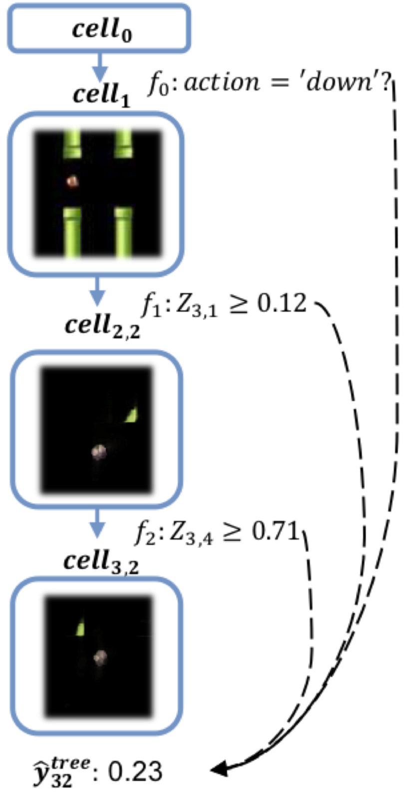
(a) Regression Tree

Illustrative Examples of Interpretable Mimic Trees

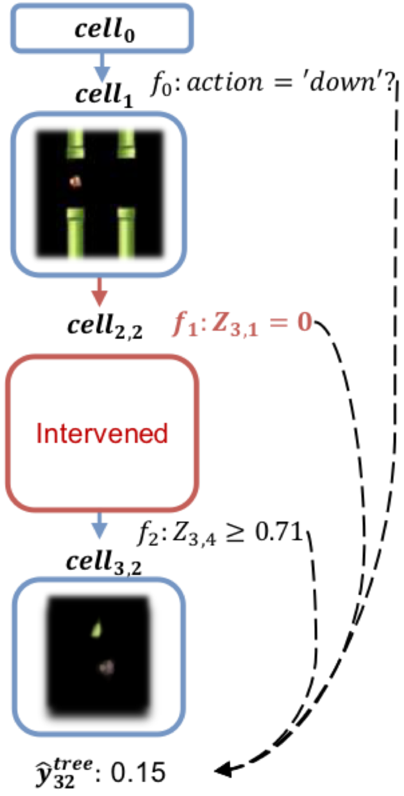
Illustrating the extracted rules:



(a) Regression Tree



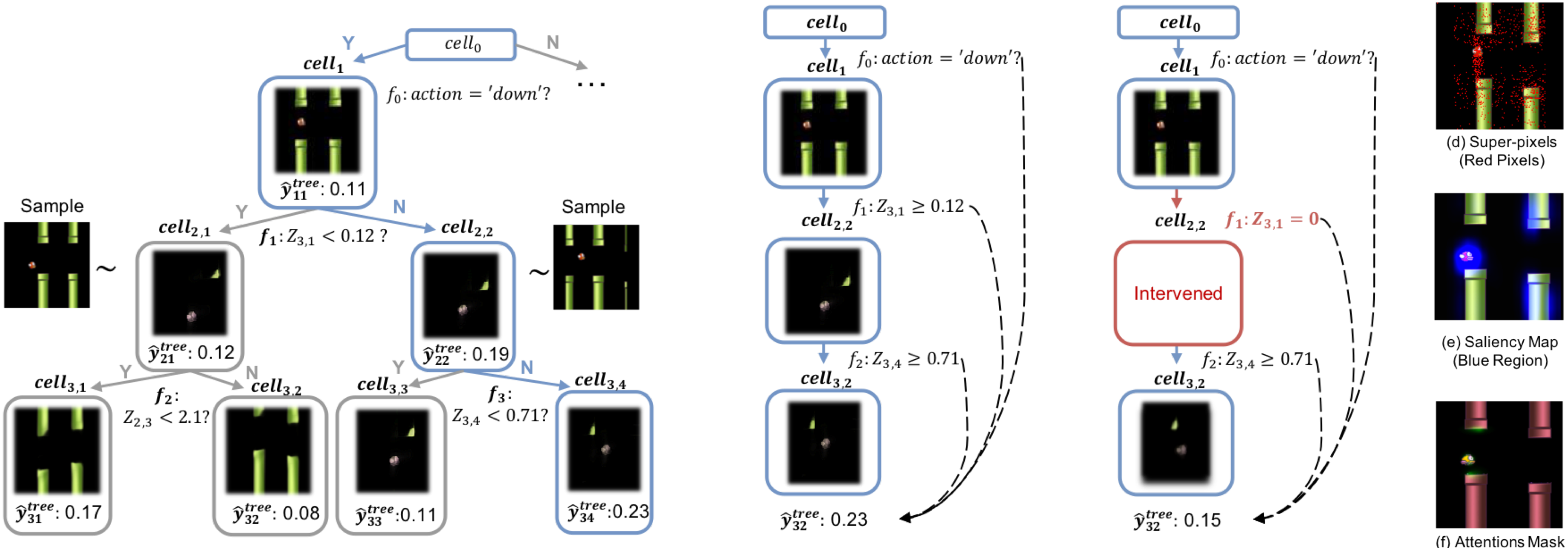
(b) Causal Relations



(c) Counterfactuals

Illustrative Examples of Interpretable Mimic Trees

Illustrating the extracted rules：



(a) Regression Tree

(b) Causal Relations

(c) Counterfactuals

(d) Super-pixels (Red Pixels)

(e) Saliency Map (Blue Region)

(f) Attentions Mask (Yellow Mask)

Conclusion

The take-home message:

- Divide the interpretation into a representation model and a decision model.
- The Information Bottleneck (IB) principle provides an effective approach for compressing input and extracting target-relevant representation.