ATRL - Project proposal

Thorben Klamt - 04.07.2024

Advanced Topics in Reinforcement Learning, Theresa Eimer, Prof. Dr. rer. nat. Marius Lindauer,
Gottfried Wilhelm Leibniz University Hannover

Backup project

Compare MAML, iMAML, evo-MAML, ES-MAML and Reptile

- Benchmark environment and configuration (Google Colab CPU servers, JAX, etc.)
- Two Benchmark environments
- Efficiency and performance in finding optimal performing HP configuration if started from (N=5) different random offsets or HP-positions in initial HP configuration

Project proposal

Leibniz Universität Hannover

Paper reference 1

- Discovering Symbolic Models from Deep Learning with Inductive Biases (2020)

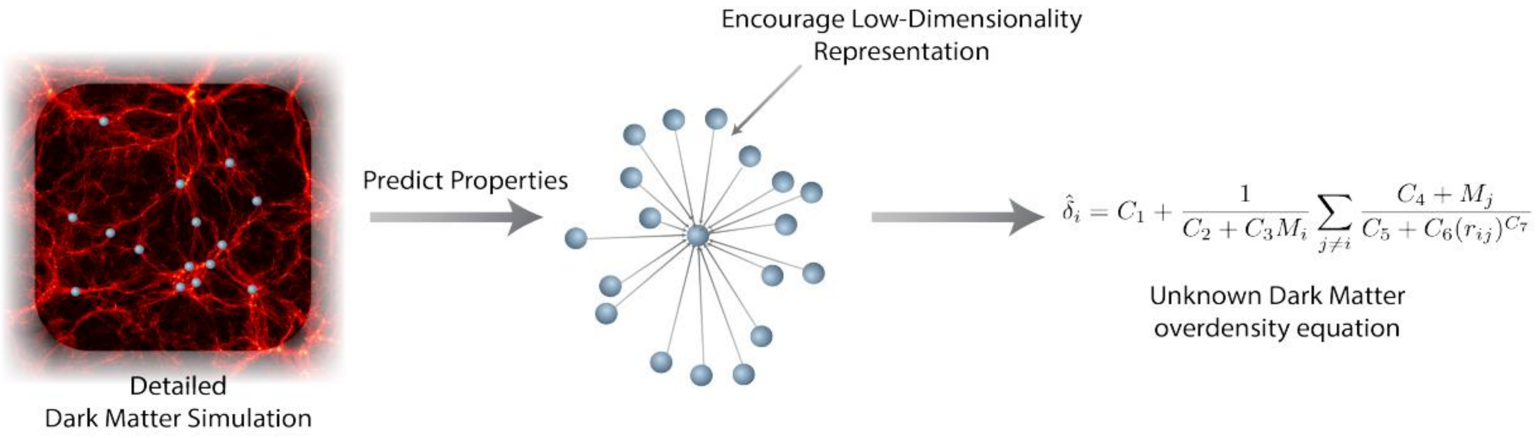

Fig:1. Cartoon depicting how physical equations are extracted from a dataset using a Deep Learning model and inductive biases in the example of Dark Matter Simulation derived overdensity equation discovery.

Paper reference 2

- Discovering Symbolic Models from Deep Learning with Inductive Biases (2020)
- Towards interpretable deep reinforcement learning with human-friendly prototypes (2023)

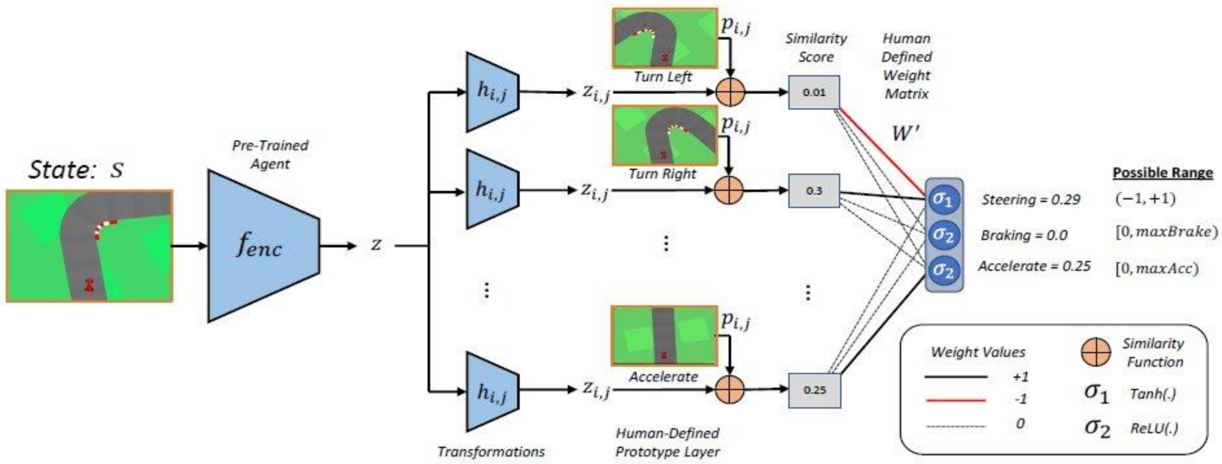

Figure 1: Prototype-Wrapper Network in Car Racing from OpenAI's gym. A state is encoded as z, transformed, and compared to human prototypes, influencing each output action.

Meta Learning meets interpretability

Meta Learning meets interpretability

Interpretable policy extraction from fully trained Deep RL agents

- Compare different methods

Meta Learning meets interpretability

Interpretable policy extraction from fully trained Deep RL agents

- Compare different methods
  - **"Conditional interpretable blocks**": pseudocode extraction, symbolic representation distillation, human-friendly prototypes
  - Two benchmark environments ((random gravity) Lunar Lander, etc.)
  - **HPO** (number of blocks, inductive bias, pseudocode length) to optimize for a two dimensional goal space: performance (normalized to baseline) versus "simplicity measure"

- Section A:
  - None
  - Deep Learning model
  - Decision tree
- Section B:
  - Human friendly prototypes
  - Symbolic representation
  - Pseudocode

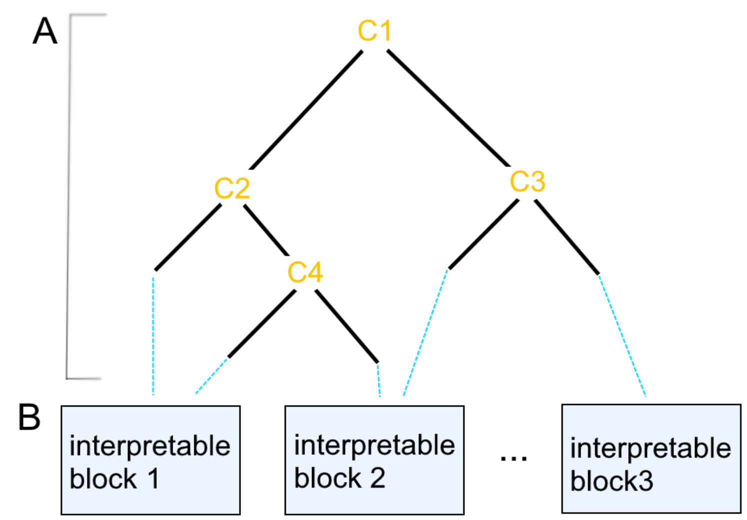

Fig:1. Abstract representation of "Conditional interpretable blocks".

Meta Learning meets interpretability

Interpretable policy extraction from fully trained Deep RL agents

- Compare different methods
  - **"Conditional interpretable blocks**": pseudocode extraction, symbolic representation distillation, human-friendly prototypes
  - Two benchmark environments ((random gravity) Lunar Lander, etc.)
  - **HPO** (number of blocks, inductive bias, pseudocode length) to optimize for a two dimensional goal space: performance (normalized to baseline) versus "simplicity measure"

- Section A:
  - None
  - Deep Learning model
  - Decision tree
- Section B:
  - Human friendly prototypes
  - Symbolic representation
  - Pseudocode

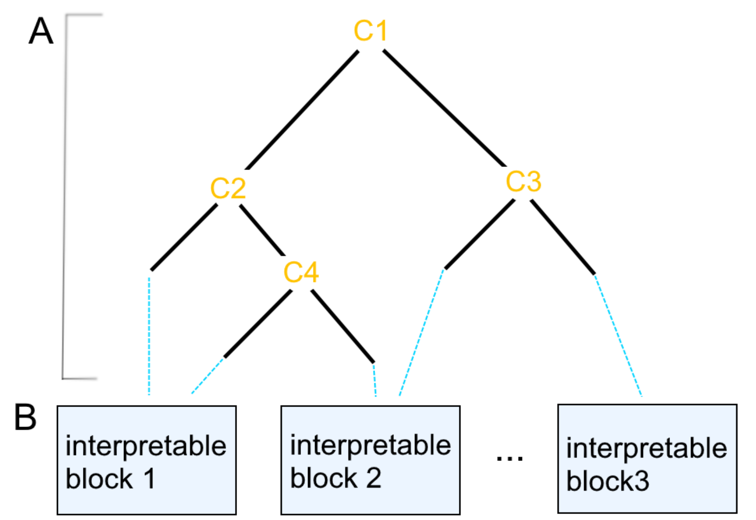

Fig:1. Abstract representation of "Conditional interpretable blocks".

Meta Learning meets interpretability

Interpretable policy extraction from fully trained Deep RL agents

- Compare different methods
  - **"Conditional interpretable blocks**": pseudocode extraction, symbolic representation distillation, human-friendly prototypes
  - Two benchmark environments ((random gravity) Lunar Lander, etc.)
  - **HPO** (number of blocks, inductive bias, pseudocode length) to optimize for a two dimensional goal space: performance (normalized to baseline) versus "simplicity measure"

- Section A:
  - None
  - Deep Learning model
  - Decision tree
- Section B:
  - Human friendly prototypes
  - Symbolic representation
  - Pseudocode

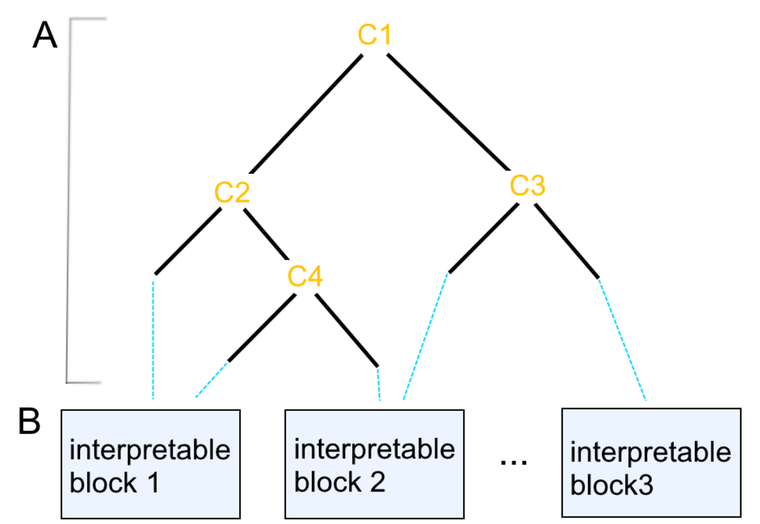

Fig:1. Abstract representation of "Conditional interpretable blocks".

Meta Learning meets interpretability

Interpretable policy extraction from fully trained Deep RL agents

- Compare different methods
  - **"Conditional interpretable blocks**": pseudocode extraction, symbolic representation distillation, human-friendly prototypes
  - Two benchmark environments ((random gravity) Lunar Lander, etc.)
  - **HPO** (number of blocks, inductive bias, pseudocode length) to optimize for a two dimensional goal space: performance (normalized to baseline) versus "simplicity measure"

- Section A:
  - None
  - Deep Learning model
  - Decision tree
- Section B:
  - Human friendly prototypes
  - Symbolic representation
  - Pseudocode

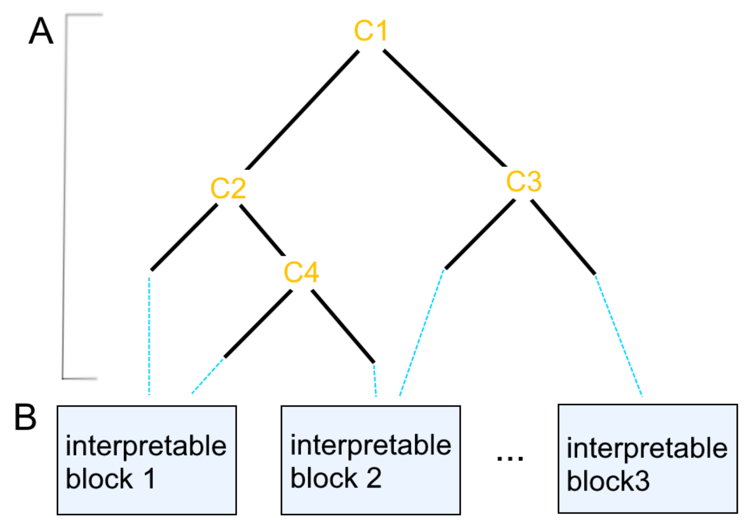

Fig:1. Abstract representation of "Conditional interpretable blocks".