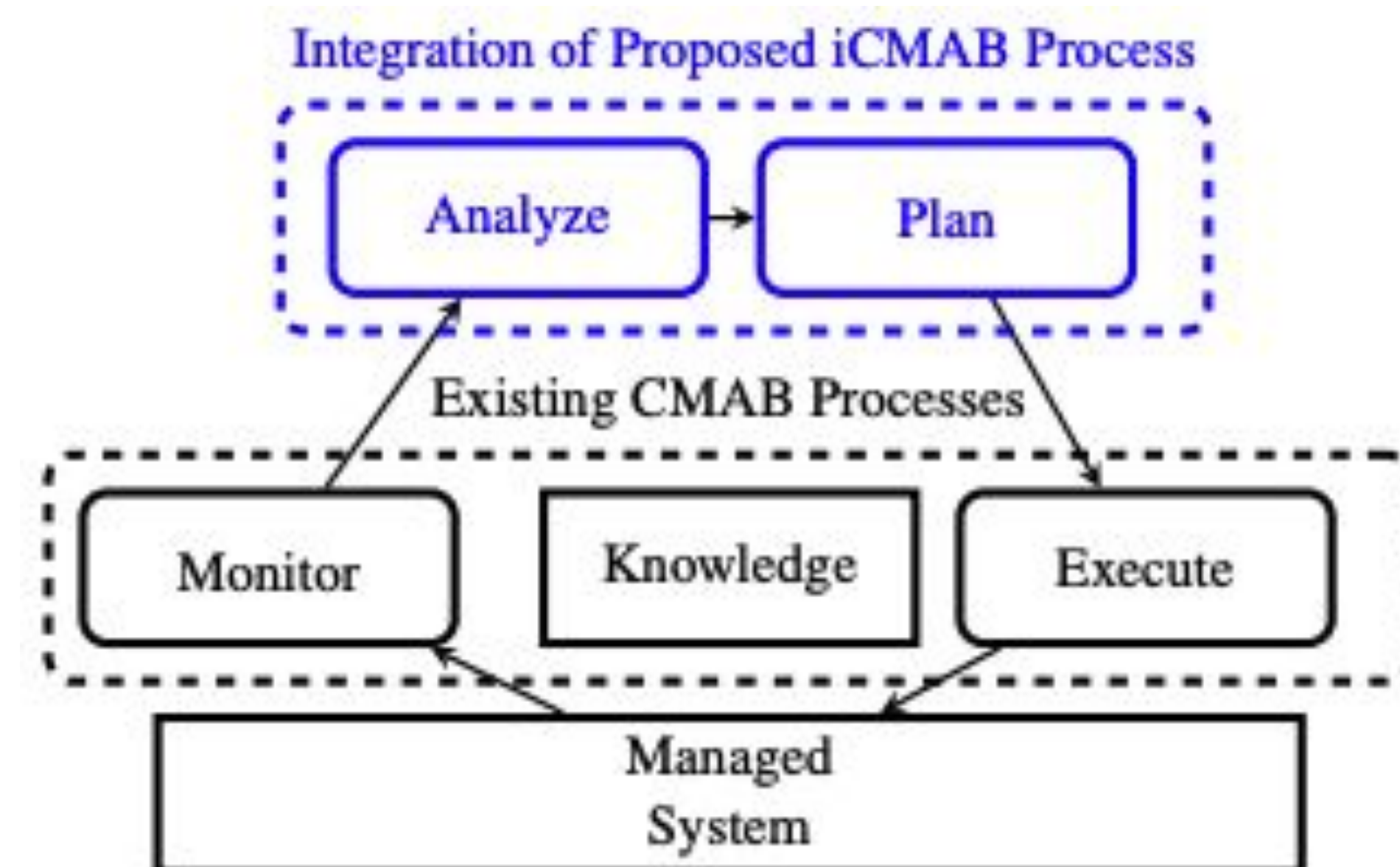


Informed Contextual Bandits to Support Decision-Making for Intelligent CPS

Travis Desell, Alex Ororbia and Daniel Krutz
Rochester Institute of Technology, Rochester, NY

Project Overview

- Multi-armed bandits (MABs) balance exploration and exploitation
- “Contextual” bandits (CMABs) incorporate context (information) in MAB process
- **Objective:** Create novel informed CMABs (iCMABs) to support CPS decision-making



Benefits

- Improve CMABs for CPS:
 - Provide prescient/missing contextual values
 - Quantify reward confidence, and identify corrupt reward values
 - Increased knowledge that is incorporated into the CPS
- Better decision-making in volatile environments
- Better adaptation to distribution shifts

Methodology

1. Reward prediction model and generative world model of contexts (dynamically adapted over time)
2. Prediction confidence/uncertainty for reward and context values
3. Online training and neuroevolution (neural architecture search) of artificial neural network (ANN) models to predict contextual and reward information

Informed Contextual Multi-Armed Bandit (iCMAB)



Current Status

- Developed a nonlinear neural regressor method for epistemic uncertainty based sample detection.
- Developed novel, online approaches to online continual learning (SOM-based)
- Designed neuroevolutionary approach for adapting RNN ensemble for time-series, multivariate regression/density estimation