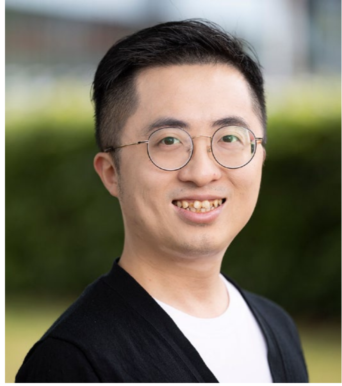


Costly and Unsafe? A good case for Reinforcement Learning

Kuan-Hsun Chen <k.h.chen@utwente.nl>, Junjie Shi <junjie.shi@tu-dortmund.de>

Assistant Professor
Computer Architecture for Embedded Systems
University of Twente, the Netherlands

technische universität dortmund

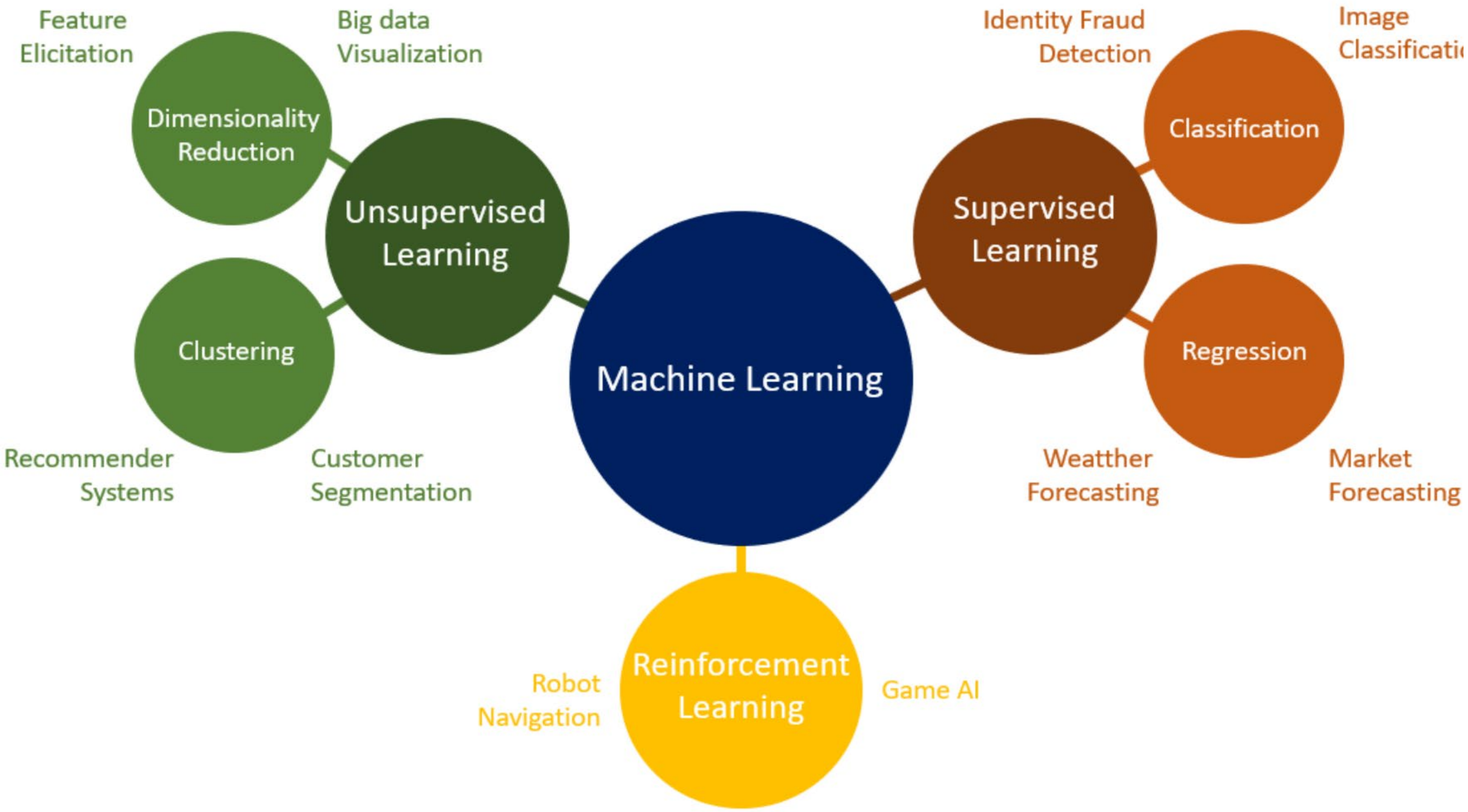
European Research Council
Established by the European Commission

UNIVERSITY OF TWENTE.

SFB 876 Providing Information by Resource-Constrained Data Analysis

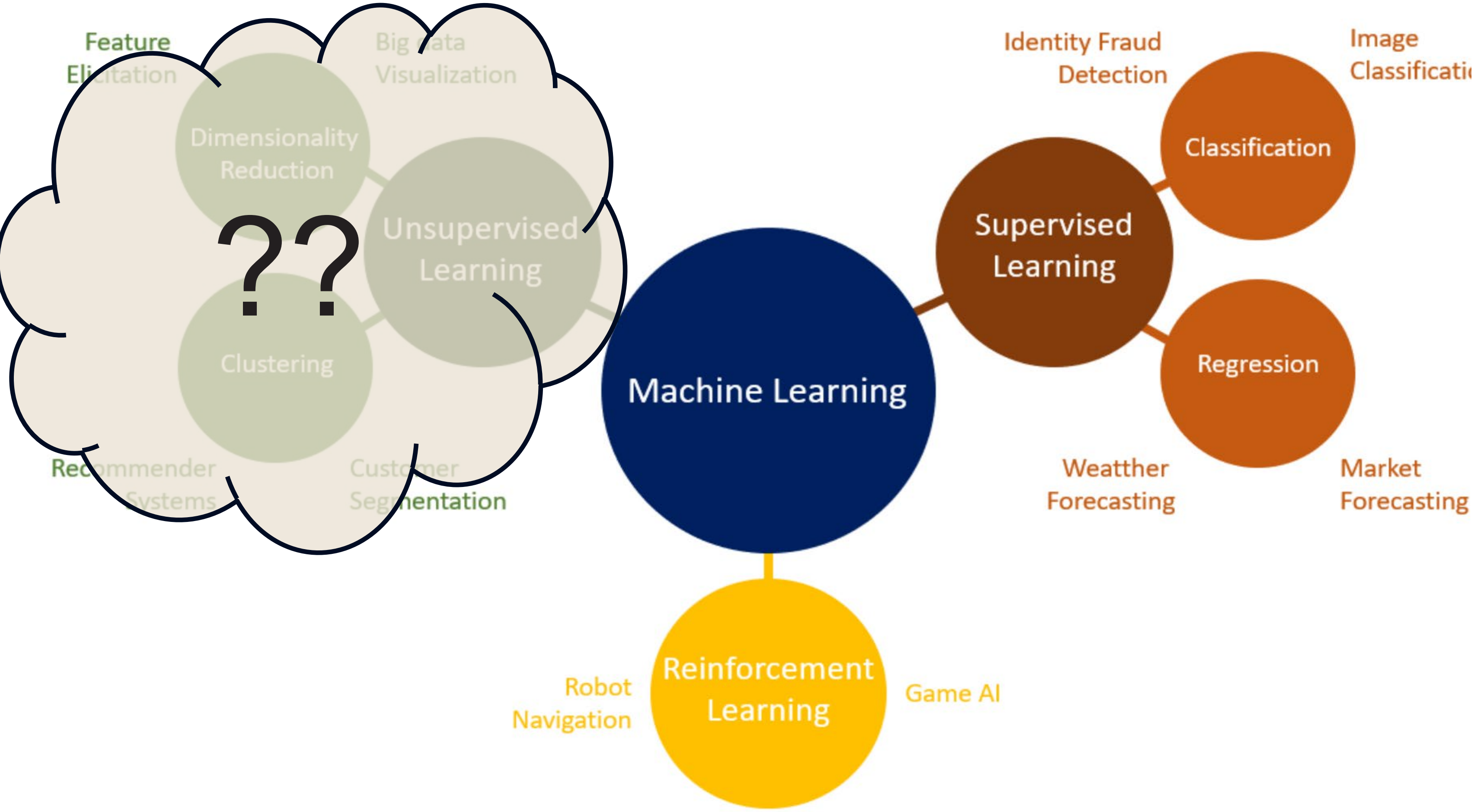
PropRT

ML for RT at the glance

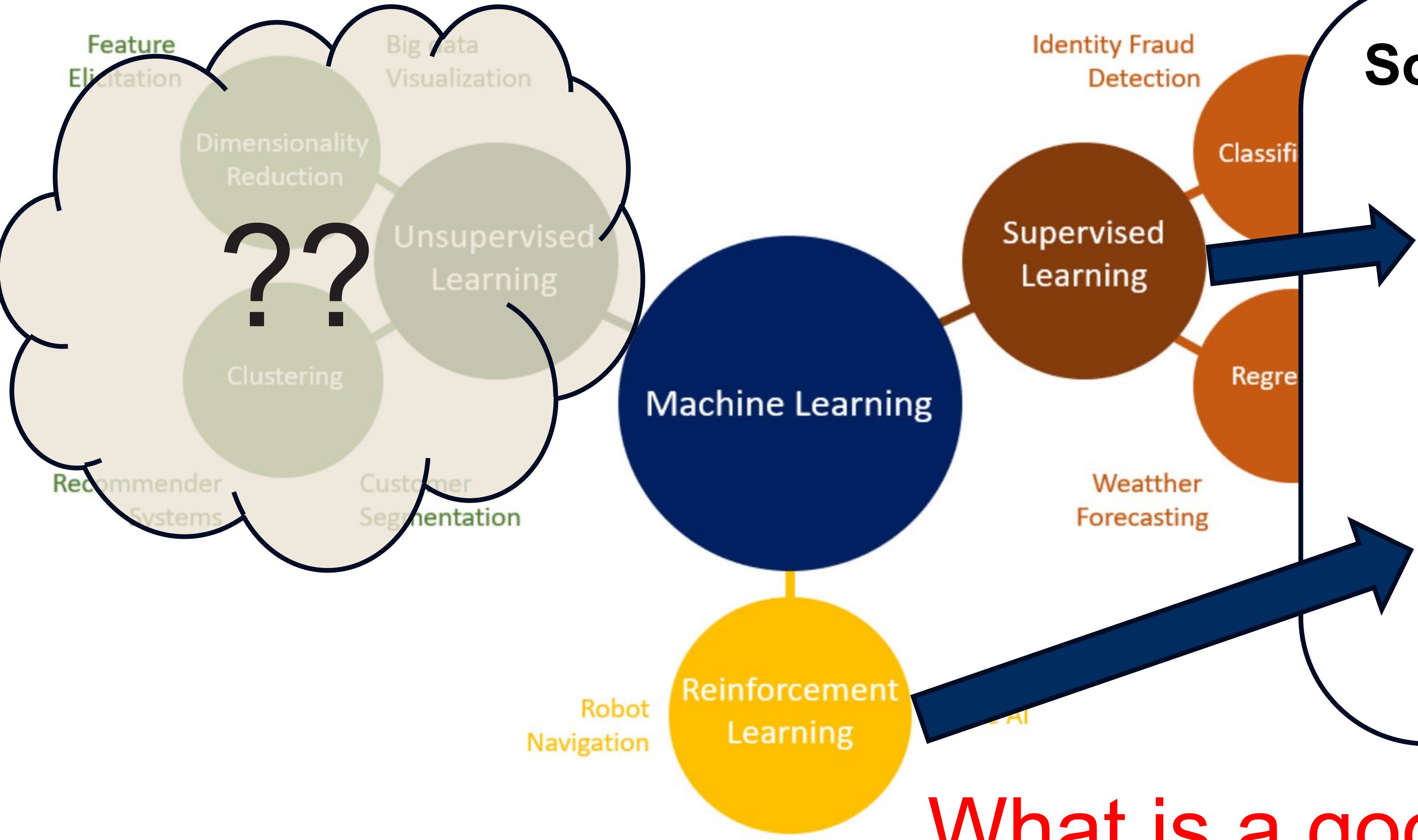


UNIVERSITY OF TWENTE.

ML for RT at the glance



ML for RT at the glance



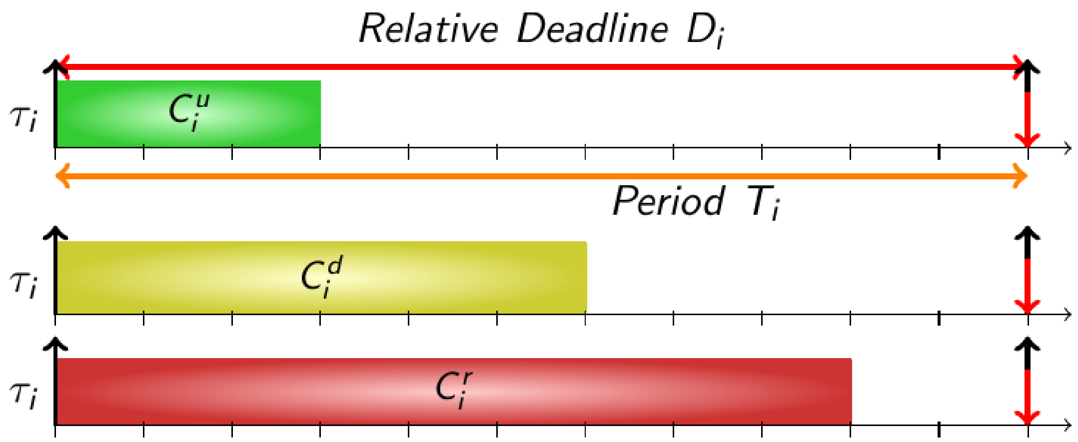
SoTA in RTS:

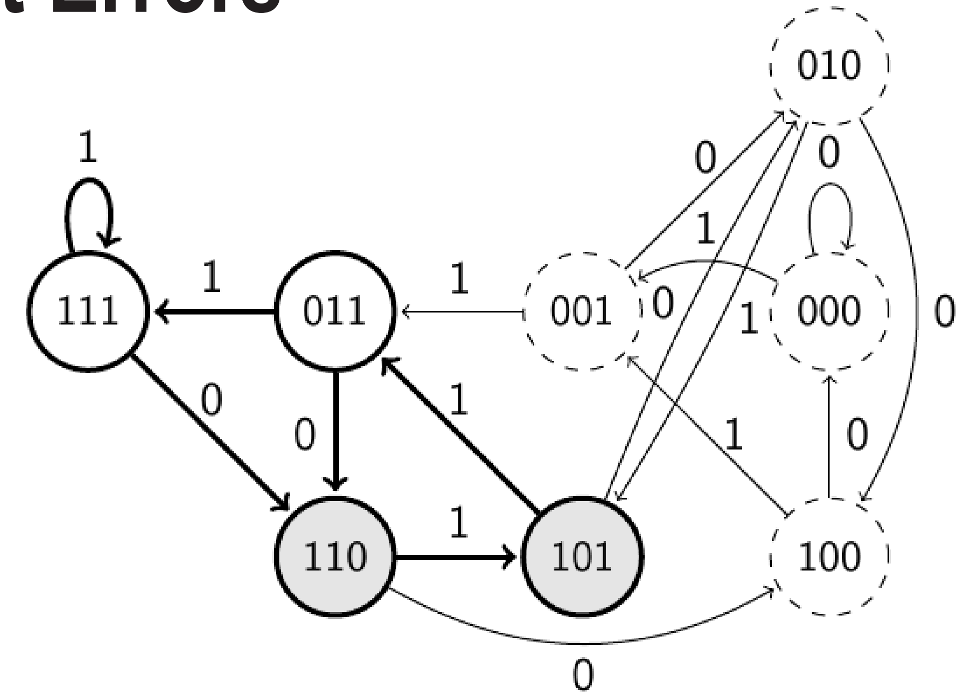
Priority Assignment for gFP [Lee et al, RTAS'21]
Labeled data is scarce

Scheduling for aperiodic tasks [Bo et al, RTAS'21]
High overhead; unsafe

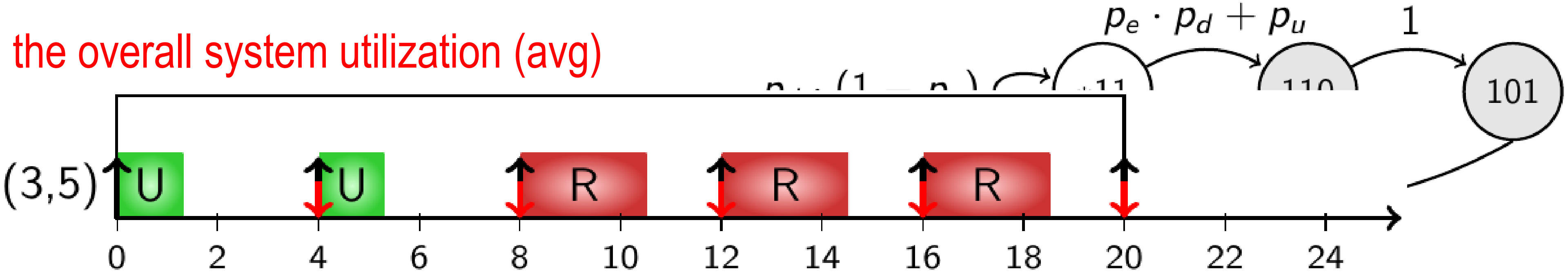
What is a good case for RL?

Safety Critical Systems against Soft Errors





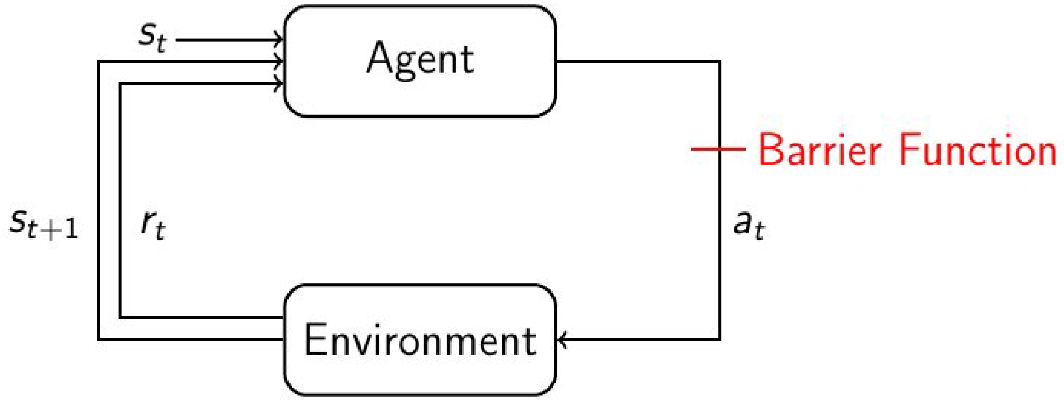
- Probability with erroneous execution p_e
- (m,k) robustness constraint
- Minimize the overall system utilization (avg)
- Formulat
- Optimiza



J. Shi, N. Ueter, J. -J. Chen and K. -H. Chen, Average Task Execution Time Minimization under (m, k) Soft Error Constraint, RTAS 2023

Reinforcement Learning Based Approach

- Markov Decision Process



- State space: environment ➔ state s_t : environment at time t
- Action space: A = {0: *unreliable*, 1: *detected*, 2: *reliable*}
- Reward function: reward r_t during transition, longer time lower reward
- Probability of state transitions: $P(s_{t+1} \mid s_t, a_t)$

J. Shi, N. Ueter, J. -J. Chen and K. -H. Chen, Average Task Execution Time Minimization under (m, k) Soft Error Constraint, RTAS 2023

Learning Formulation, Reward, and Barrier Function

- Maximized the cumulative reward (i.e., execution time)
- Example when (m = 3, k = 5):

$$
\begin{array}{r}
\\
\text{Correctness} \\
\text{Execution mode} \\
\text{Expected execution time} \\
\text{Real execution time}
\end{array}
\begin{array}{c}
\begin{array}{ccccc} J^1 & J^2 & J^3 & J^4 & J^5 \end{array} \\
\left(\begin{array}{ccccc}
1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 2 \\
C^d & C^u & C^d & C^d & C^r \\
C^d & C^u & C^d & C^d & C^r
\end{array}\right)
\end{array}
\longrightarrow
\begin{array}{c}
\begin{array}{ccccc} J^2 & J^3 & J^4 & J^5 & J^6 \end{array} \\
\left(\begin{array}{ccccc}
0 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 2 & 1 \\
C^u & C^d & C^d & C^r & C^d \\
C^u & C^d & C^d & C^r & \boxed{C^d + C^r}
\end{array}\right)
\end{array}
$$

$$s_t \xrightarrow[a_t = 1]{} s_{t+1}$$

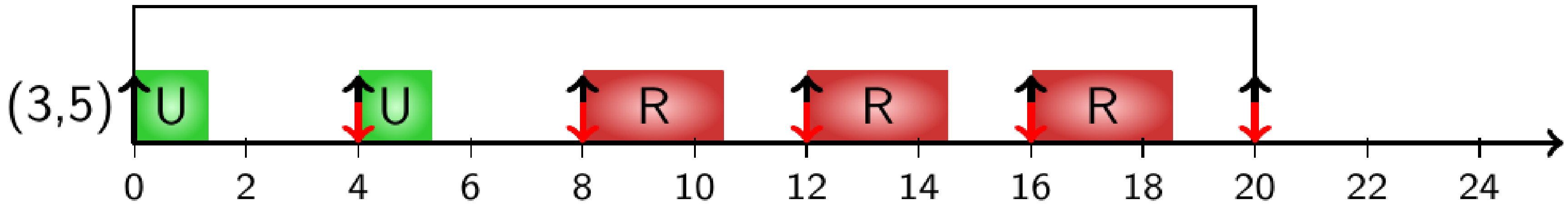
Barrier function to <u>nudge</u> the action of <u>critical state</u>

1. Reliable mode: no action
2. Detected mode: follow up reliable mode if error is detected
3. Unreliable mode: forbidden, extreme large negative reward

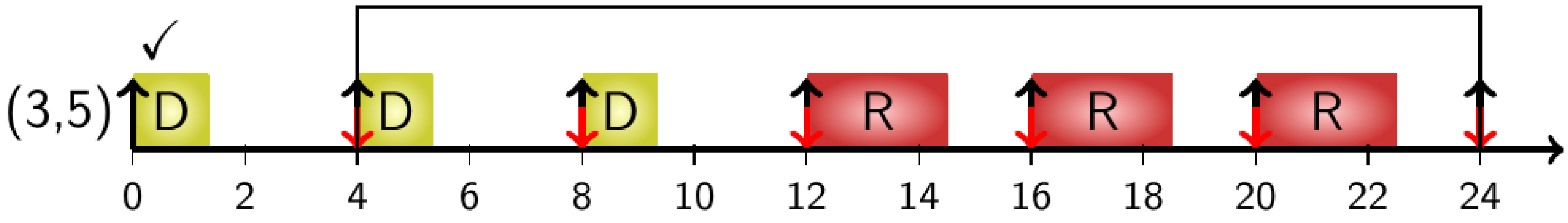
J. Shi, N. Ueter, J. -J. Chen and K. -H. Chen, Average Task Execution Time Minimization under (m, k) Soft Error Constraint, RTAS 2023

Evaluation Setup

- Optimized mapping strategy (OPT)
- DQN with 10-layer neural network (RL)
- Static approach (STA) with R-Pattern [Chen, LCTES'16]

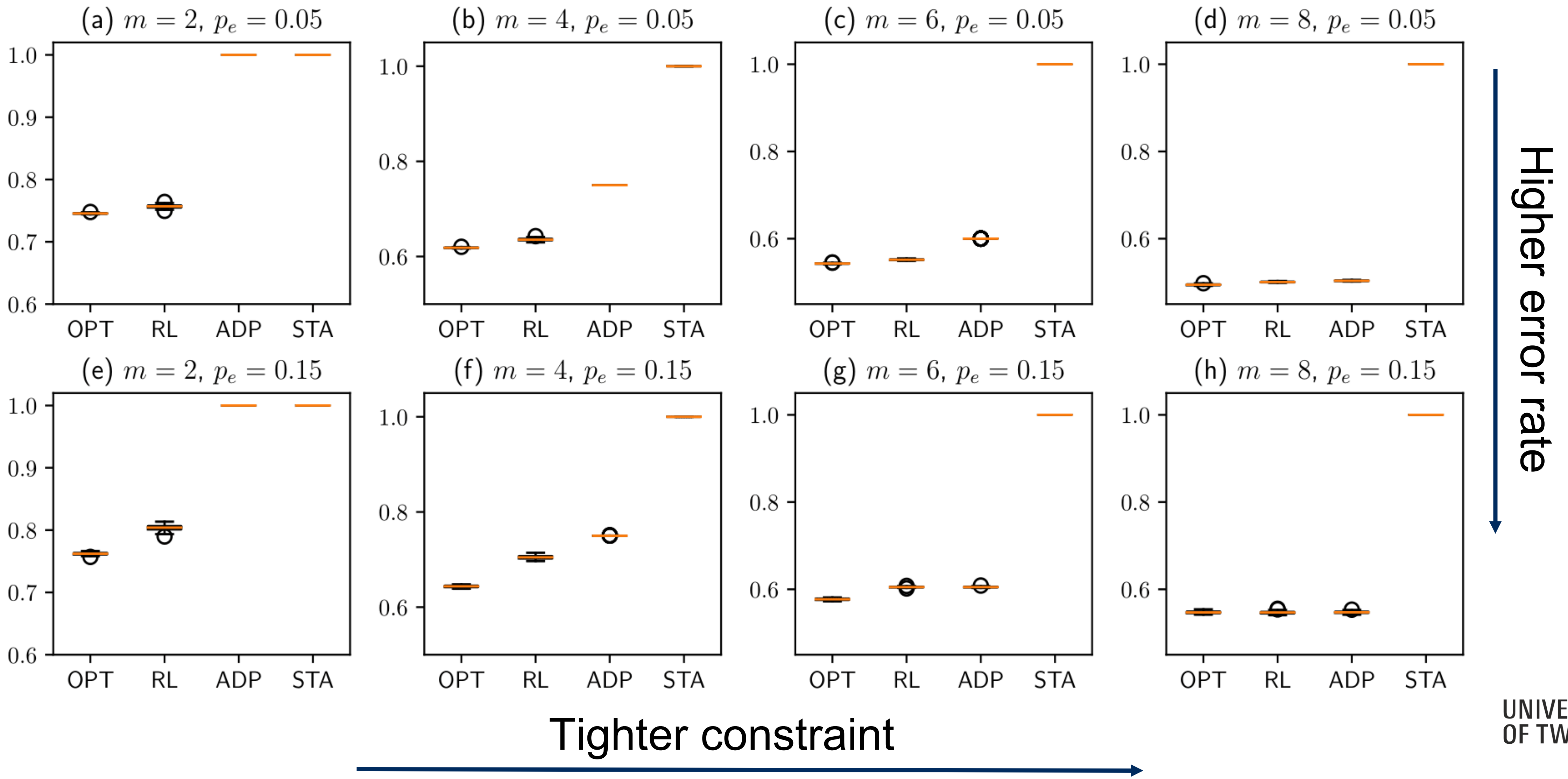


- Adaptive approach (ADP) with R-pattern [Chen, LCTES'16]



J. Shi, N. Ueter, J. -J. Chen and K. -H. Chen, Average Task Execution Time Minimization under (m, k) Soft Error Constraint, RTAS 2023

Evaluation Results: k = 10 (the lower the better)

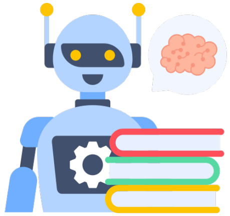


UNIVERSITY OF TWENTE.

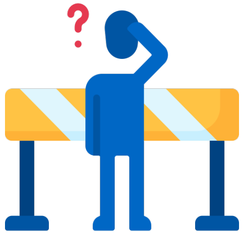
Take-home Message

RL is attractive for handling dynamic cases

Limited states can make RL efficient



Barrier function is the key for safety



UNIVERSITY
OF TWENTE.

Thank you! Questions?