# Machine (or Reinforcement): Learning to assist vessel docking in extreme environments

Bachelor defense

June 18, 2020

**Anne-Charlotte Poulsen**

SDU Robotics
The Maersk Mc-Kinney Moller Institute
University of Southern Denmark

# Agenda



- Quick introduction to the problem
- Reinforcement Learning
- Simplified model
- Results & Discussion
- Conclusion & Future work

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma’s current solution?
- The objective
- The approach

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma's current solution?
- The objective
- The approach

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma's current solution?
- The objective
- The approach

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma's current solution?
- The objective
- The approach

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma's current solution?
- The objective
- The approach

Overview of the problem

- Vessel stabilizer
- Why is a vessel stabilizer needed?
- What is Dacoma's current solution?
- The objective
- The approach

Overview:

- Classic Reinforcement Learning & Terminology
- The Approach
- Issues and Shortcomings

Terminology

- Agent
- Environment
- States, Statespace, Actions & Actionspace
- Reward and Reward function
- Policy

Classic Reinforcement Learning (CRL)

- Markov Decision Process (MDP)
- Problem Review
- Why CCRL does not fit on this problem

Concepts:

- Policy Gradient Methods
- Neural Networks
- Optimizer

Implementation specifics:

- Finite differences
- Reward & Reward Function

Implementation specifics:

- Finite differences
- Reward & Reward Function

Implementation specifics:

- Finite differences
- Reward & Reward Function

Matematiske og grafiske metoder til syntese af lineære tidsinvariante systemer:[1]

- diskret og kontinuert tilstandsbeskrivelse
- analyse i tid og frekvens
- stabilitet, reguleringshastighed, følsomhed og fejl
- digitale PI, PID, LEAD og LAG regulatorer (serieregulatorer)
- tilstandsregulering, pole-placement og tilstands-estimering (observer)
- optimal regulering (least squares) og optimal tilstands-estimation (Kalman-filter)

**Færdigheder**:
Efter gennemførelse af kurset kan den succesfulde studerende:

- kunne analysere, dimensionere og implementere såvel kontinuert som tidsdiskret regulering af lineære tidsinvariante og stokastiske systemer

**Kompetencer**:
Efter gennemførelse af kurset kan den succesfulde studerende:

- anvende og implementere klassiske og moderne reguleringsteknikker for at kunne styre og regulere en robot hurtig og præcist

[1] Based on https://fagbesk.sam.sdu.dk/?fag_id=39673