“What does vehicle automation with Meaningful Human Control mean in practice?”

Dr. Ir. S.C. Calvert – STAD/MHC College Tour – August 2020

Meaningful Human Control
over automated driving systems

TU Delft

Introduction

AUTOS | News

Study: Autonomous vehicles won't make roads completely safe

Tom Krisher
The Associated Press Staff
Contact

Published Thursday, June 4, 2020 12:18PM EDT
Last Updated Thursday, June 4, 2020 3:33PM EDT

MOST-READ

This map shows why it would be 'terrifying' to open Canada-U.S. bord

The Washington Post
Democracy Dies in Darkness
Get 1 year for $29

Technology

What self-driving cars can't recognize may be a matter of life and death

Engineers are racing to program artificial intelligence to recognize different scenarios that human drivers know inherently

BUSINESS CULTURE GEAR IDEAS SCIENCE MORE

Stop Saying Driverless Cars Will Help Old People

And maybe start including them in research instead of just assuming we know what they want.

The Economist
Today Weekly edition Menu

Leaders
Oct 10th 2019 edition

Autonomous vehicles

Driverless cars are stuck in a jam

Blame Silicon Valley hype—and the limits of AI

Outline

- Recap on what MHC is
- Application and operationalisation
- Core components
- Application to Truck Platooning
- Quantitative operational test case

Meaningful Human Control

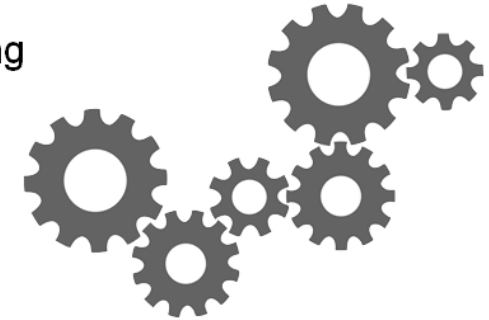
Meaningful Human Control
over automated driving systems

- Philosophy & Ethics
- Behaviour & Psychology
- Traffic & Engineering

TU Delft

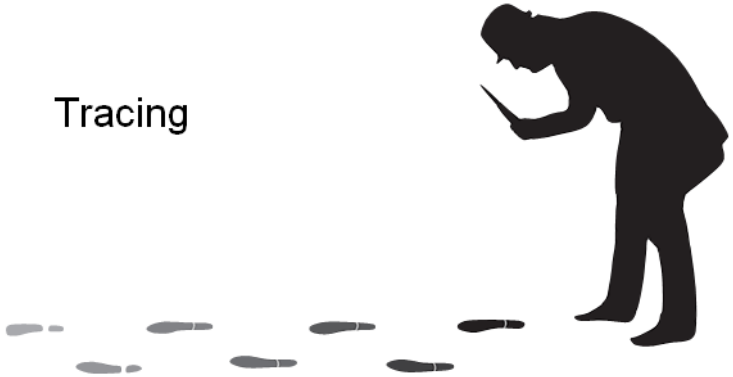
MHC conditions

Tracking

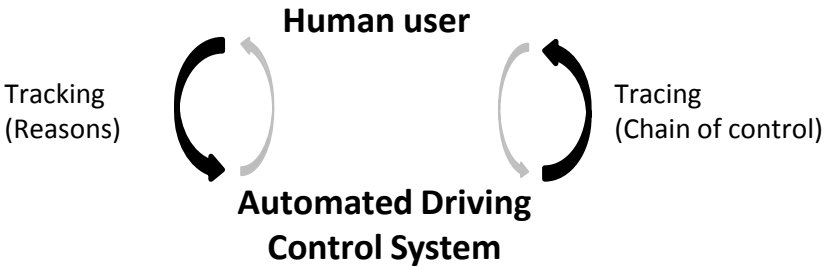


The system (human operators, operated devices, infrastructures…) should be able to co-vary its behavior with the relevant reasons of the relevant human agent(s) for carrying out X or omitting X

Tracing



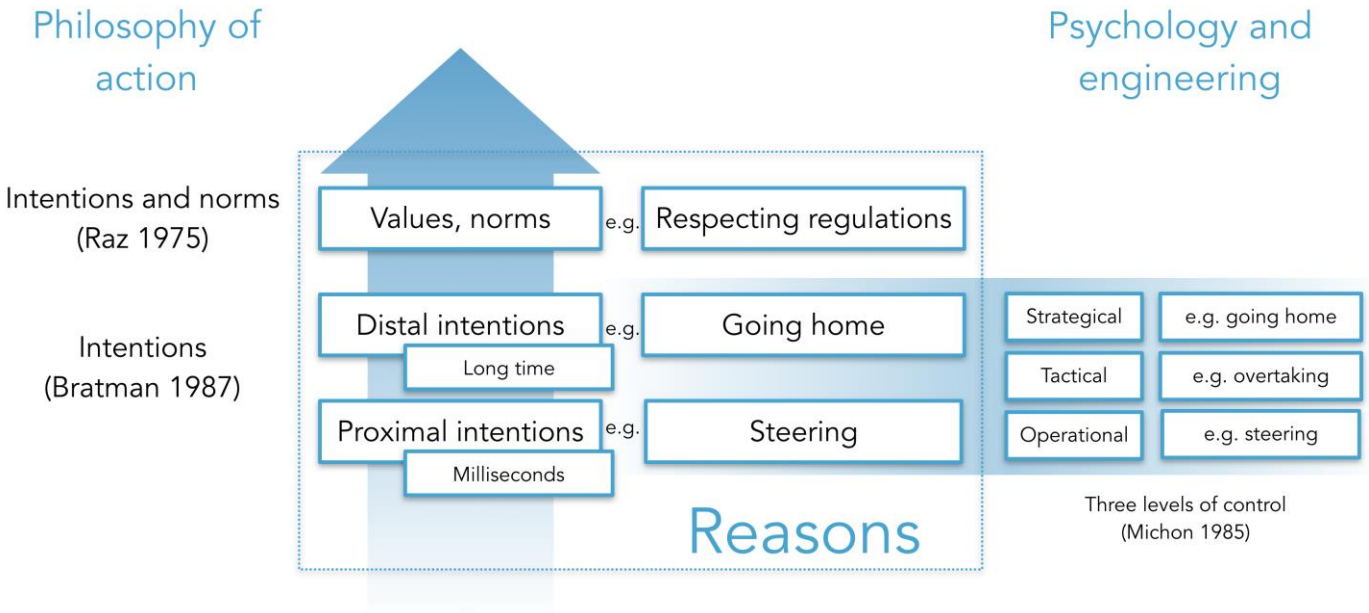
There is at least one human agent in the system design history or use context who can appreciate the capabilities of the system and her own role as target of potential moral consequences for the system's behaviour



TU Delft

Control:
Connecting philosophy & engineering

Expanding a classic theory of control with intuitions from philosophy of action



G. Mecacci & F. Santoni de Sio (2019). "Meaningful human control as reason-responsiveness: the case of dual-mode vehicles". Ethics and Information Technology.

Poll: How well do you think human ability and morals are considered in current AD-systems?

A. Very well, almost completely
B. The main aspects are included
C. Scarcely, only some aspects are considered
D. Not at all

Challenges towards application

- Very abstract concept
  (in contrast to physical operational control)
- Complex ethical and behavioural interactions
- Collectively subjective

- Operationalisation allows application in practice

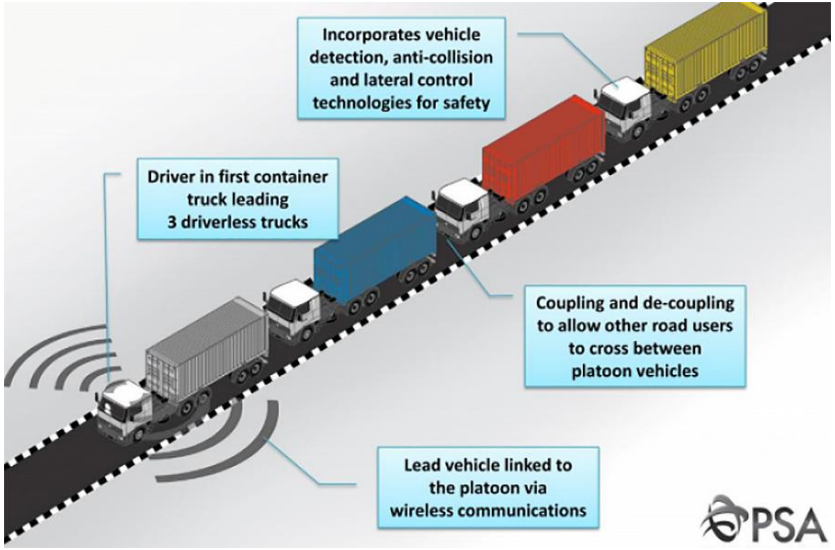
Challenges towards application

- Very abstract concept
  (in contrast to physical operational control)
- Complex ethical and behavioural interactions
- Collectively subjective

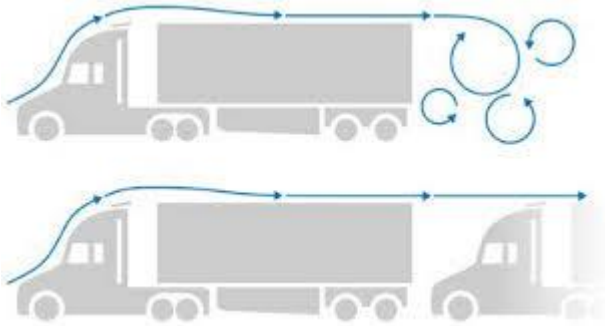
- Operationalisation allows application in practice

1. (Definition of core components)
2. (Simulation modelling framework)
3. Conceptual application of MHC
4. Quantification of MHC conditions

Case study: Truck platooning



- Cooperative
- (Partially) automated

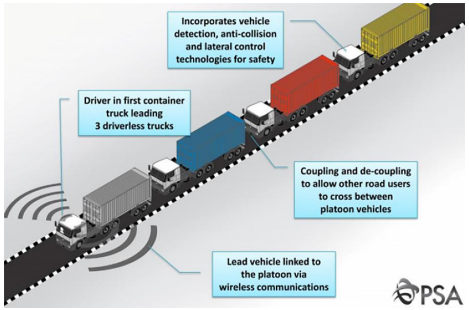


Source: NREL

- Fuel savings
- Lower emissions
- Lower labour costs?
- Traffic efficiency?

TP control (physical & practical)

- Operational control of **platoon**:
  - Front driver & (C)ACC-system
  - Human control: Front driver
- Operational control of **following trucks**:
  - (C)ACC-system
  - Following driver: monitors truck and system
  - Human control: Front driver !?
- Responsibility:
  - Lies with each individual driver (in principle!)
- ‘Effective’ control is potentially confused or incomplete!



S. C. Calvert, G. Mecacci, D. D. Heikoop, & F. Santoni de Sio (2018). “Full Platoon Control in Truck Platooning: A Meaningful Human Control Perspective”. IEEE ITSC conference, November 4-7, 2018, Maui, USA.

TP control: example

- Situation:
  - TP approaches workzone
  - CACC-system doesn't detect or makes takeover request
  - Front driver is distracted and reacts late
  - Front driver disengages the system by performing an emergency manoeuvre without crashing
  - The following truck drivers cannot react in time and collide with workzone and each other

S. C. Calvert, G. Mecacci, D. D. Heikoop, & F. Santoni de Sio (2018). "Full Platoon Control in Truck Platooning: A Meaningful Human Control Perspective". IEEE ITSC conference, November 4-7, 2018, Maui, USA.

TP control: example

- Situation:
  - TP approaches workzone
  - CACC-system doesn't detect or makes takeover request
  - Front driver is distracted and reacts late
  - Front driver disengages the system by performing an emergency manoeuvre without crashing
  - The following truck drivers cannot react in time and collide with workzone and each other
- Problem
  - The following drivers are held responsible because they were meant to monitor their own situation of their own trucks

S. C. Calvert, G. Mecacci, D. D. Heikoop, & F. Santoni de Sio (2018). "Full Platoon Control in Truck Platooning: A Meaningful Human Control Perspective". IEEE ITSC conference, November 4-7, 2018, Maui, USA.

Applying MHC to examples

- Tracking (‘system’ compliance to…):
  - Yes, platoon is en-route and compliant
  - Although in an instance, the system cannot perform an emergency manoeuvre, but that is situational rather than by intention.
- Tracing (actor capable to control system…):
  - Front driver: Not met: Performed delayed emergency manoeuvre – also puts too great demands on system
  - Following drivers: Not met: unrealistic transition of control demand
  - ADS designer: No MHC, outside ODD (hence TO-request made)

- Both system design and driver performance translate to a lack of MHC (even before an accident) for the case

TP control and challenges

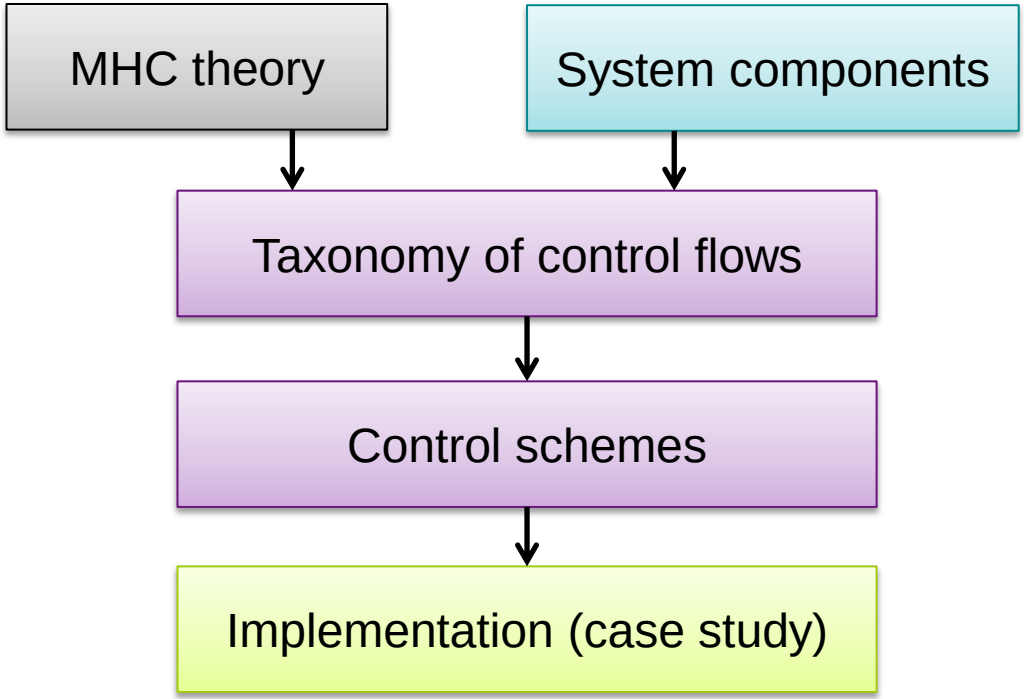
- Operational control and responsibility not aligned
- Misbalance in:

TP control and challenges

- Operational control and responsibility not aligned

- Misbalance in:
  - Operational control vs ‘effective human’ control
  - Behavioural capabilities of drivers (cognitive)
  - Ethical issues (demands, accountability and responsibility)

Operationalisation of MHC

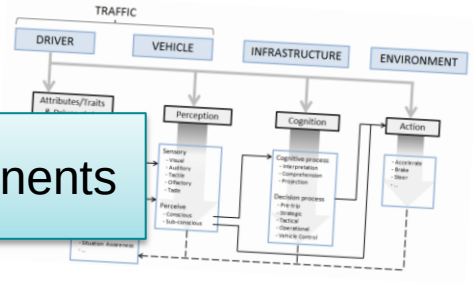


Operationalisation of MHC

Tracking
Tracing

MHC theory

System components



Taxonomy of control flows

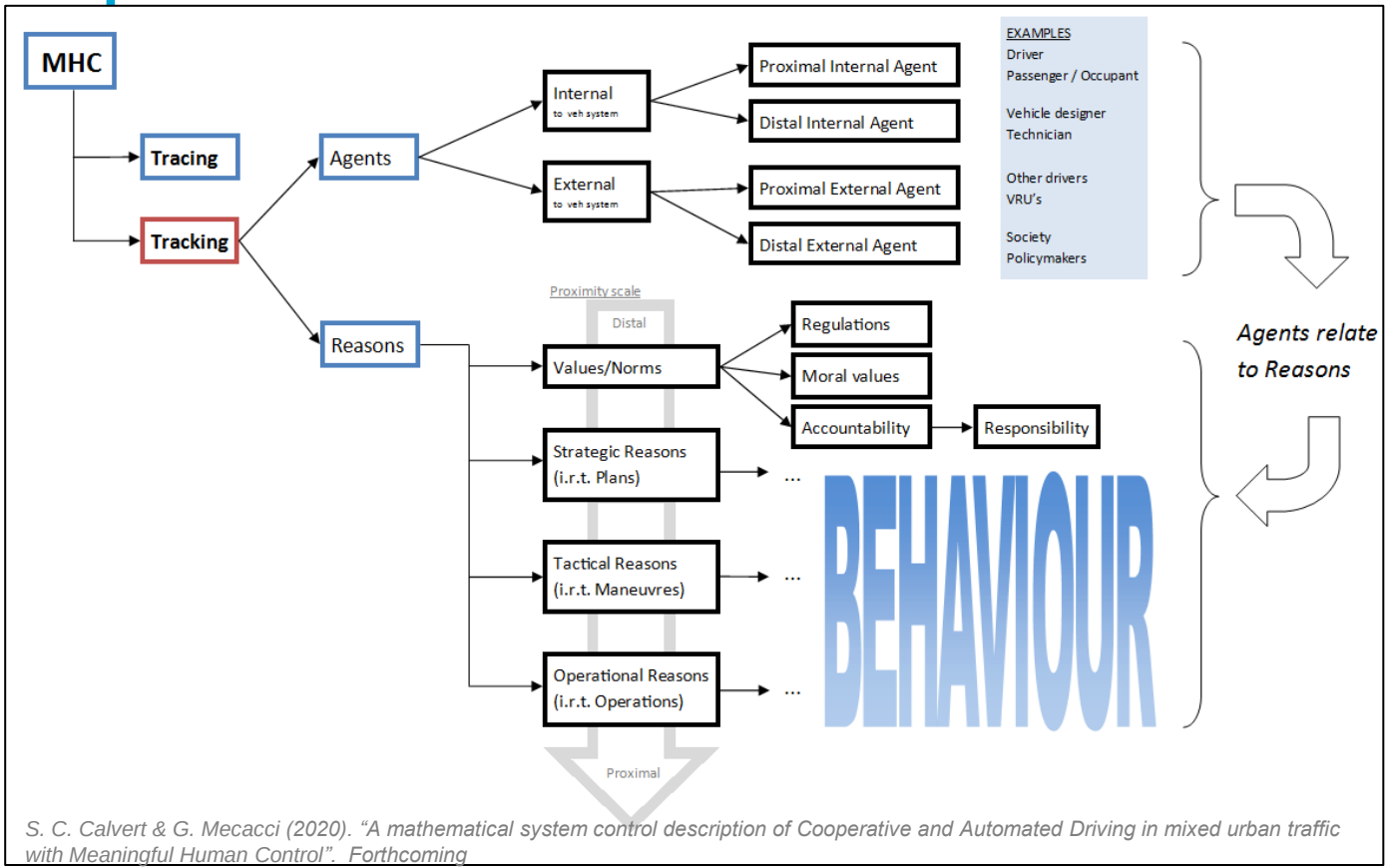
Simplified!

Control schemes

Based on a learning system!

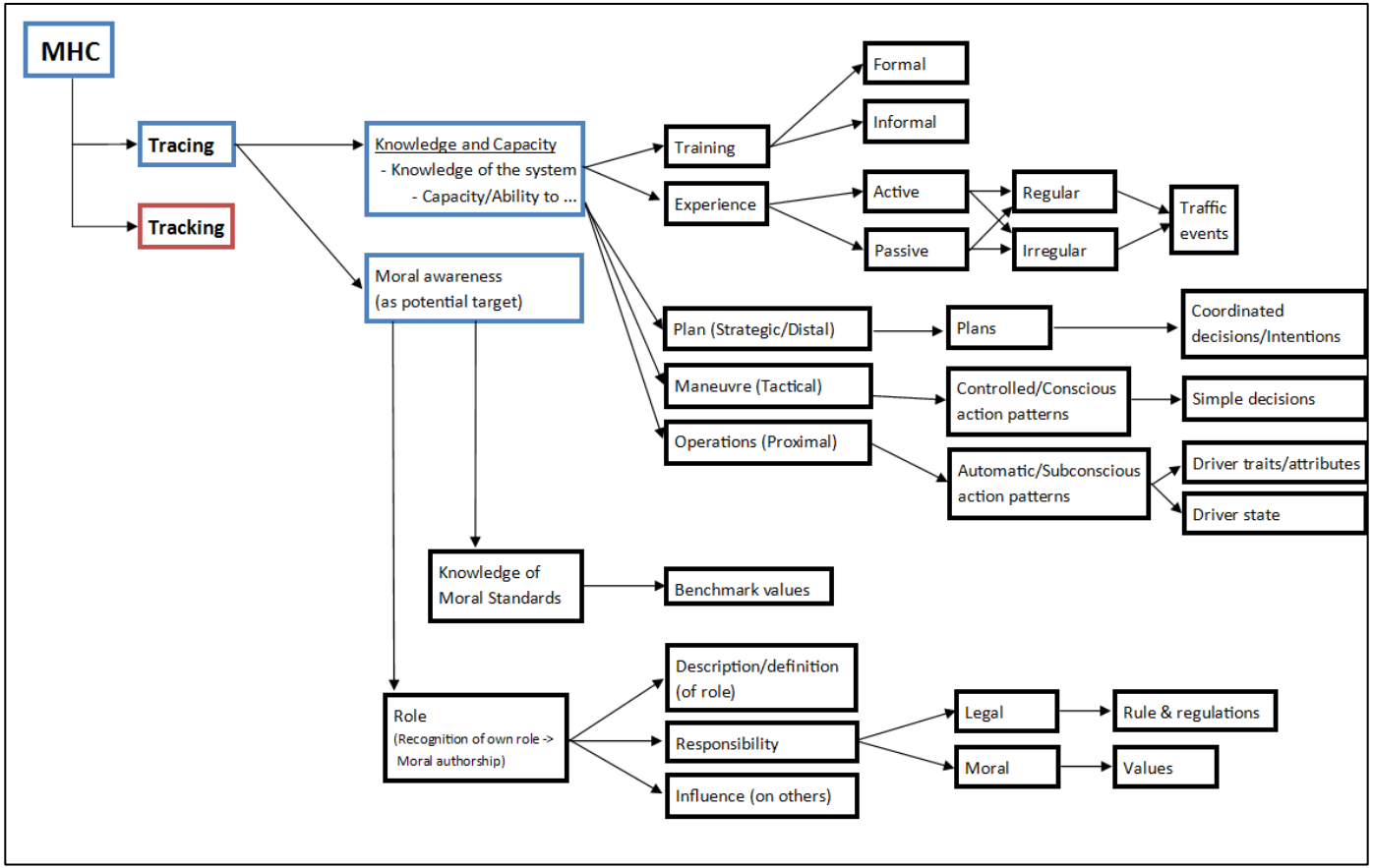
Implementation (case study)

Operationalisation of MHC

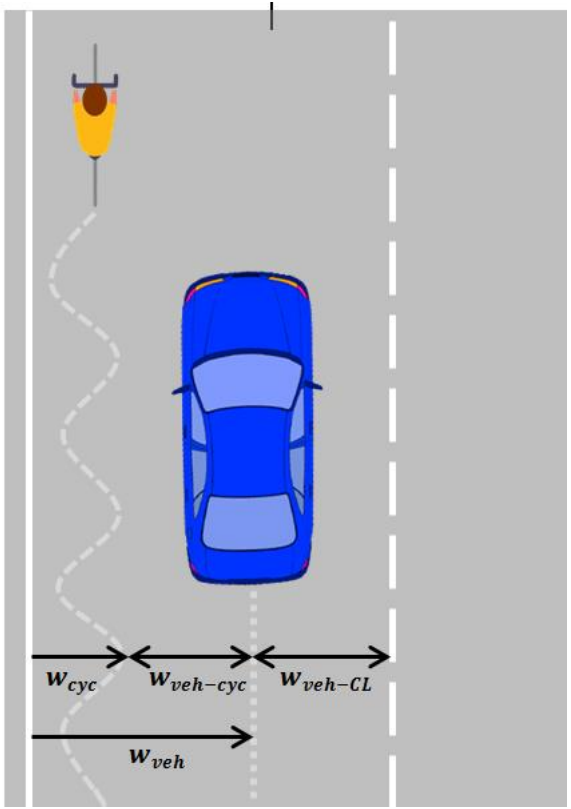


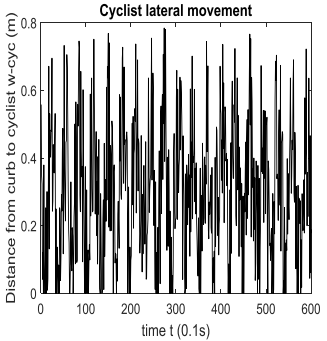
S. C. Calvert & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

Operationalisation of MHC



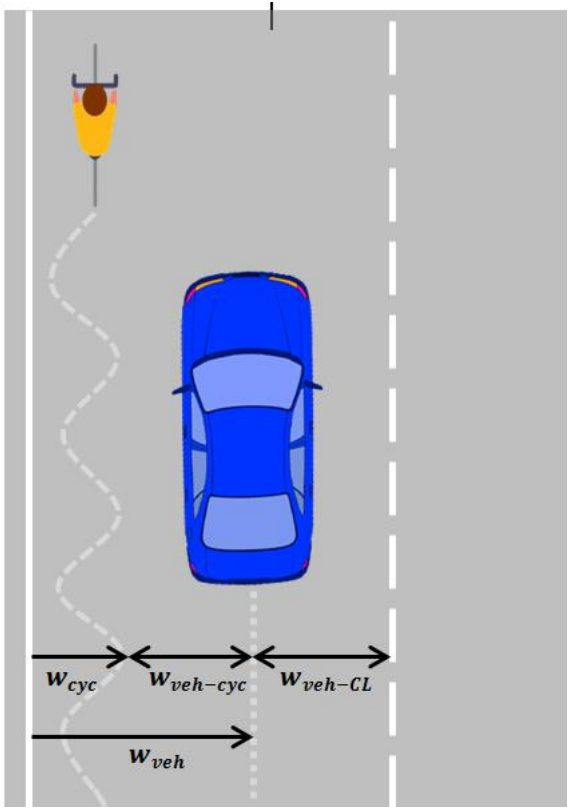
Operationalisation of MHC





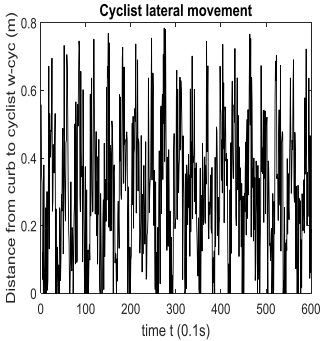
S. C. Calvert, & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

Operationalisation of MHC



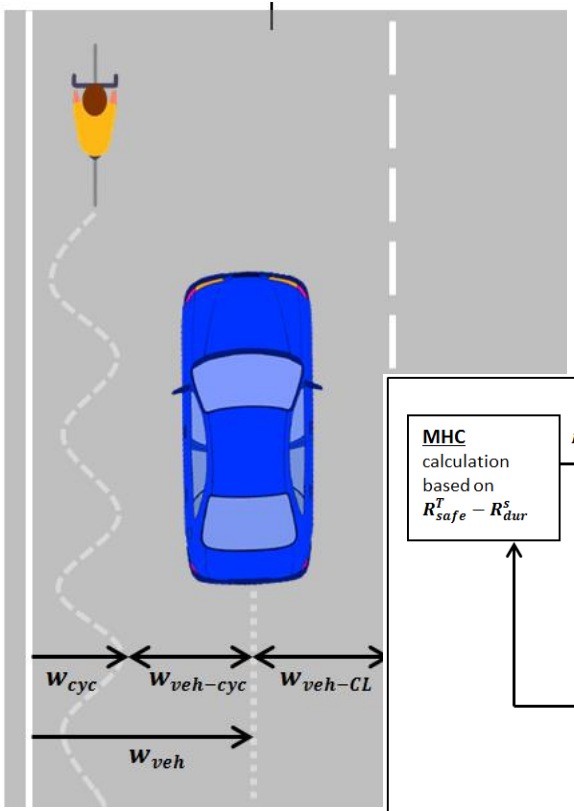
Reaction to ‘reasons’ of:

- Duration (strategic)
- Safety (tactical)



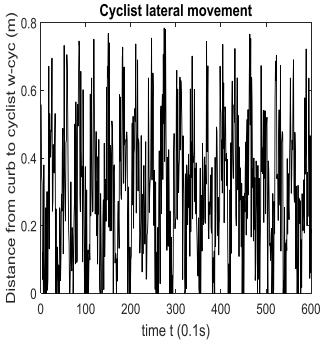
S. C. Calvert, & G. Mecacci (2020). “A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control”. Forthcoming

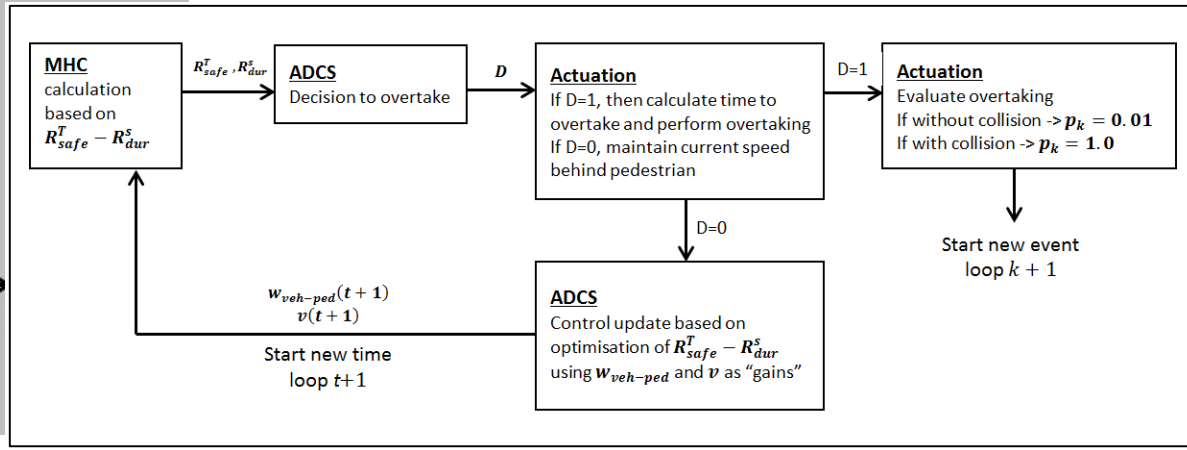
Operationalisation of MHC



Reaction to 'reasons' of:

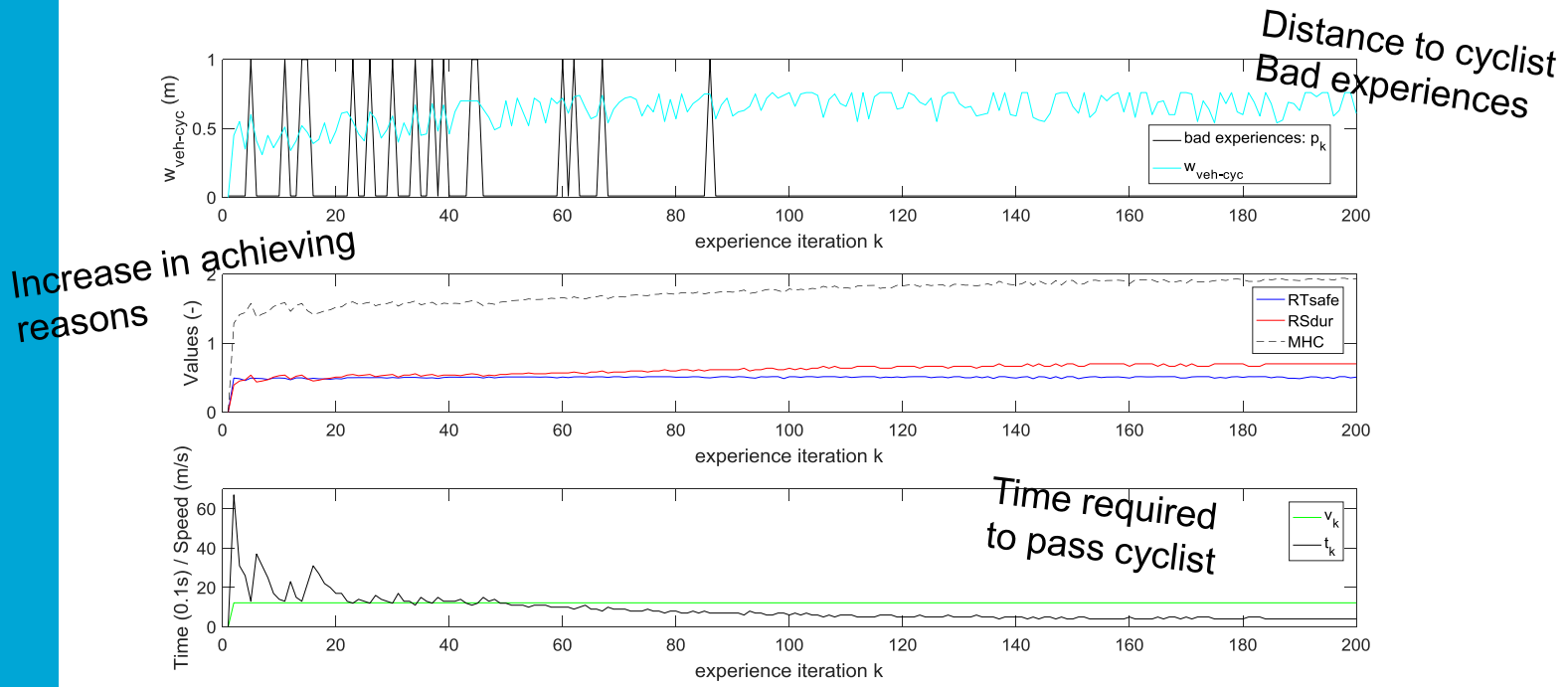
- Duration (strategic)
- Safety (tactical)





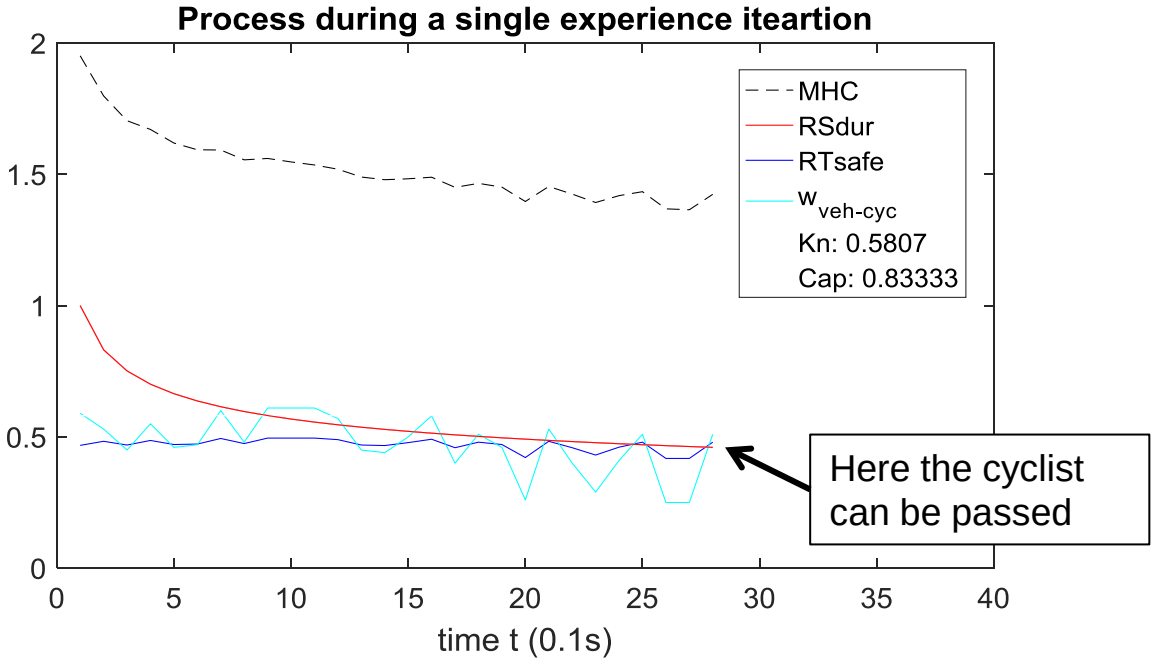
S. C. Calvert, & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

Operationalisation of MHC



S. C. Calvert, K. Ampountolas & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

Operationalisation of MHC



S. C. Calvert, K. Ampountolas & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

Operationalisation of MHC

Why is operationalisation important?

- Makes abstract concept applicable in practice
- Demonstrates ways that MHC can be considered in vehicle and infrastructure design
- Demonstrates an approach to evaluate the extent of MHC
- Demonstrates potential policy influence on MHC

TU Delft

Question: Where do you see potential applications of MHC in your work practice or domain?

Get involved!

Not the end of the story

- Proper AV-systems require ethical and behavioural considerations!
- How does your AV-system consider this?
- Are you willing to investigate options to optimise you're AV-system
- *We now have the building blocks, so let's build!*

Questions

Simeon C. Calvert, PhD
s.c.calvert@tudelft.nl

TU Delft