IEEE LARS 2018 – 15th Latin American Robotics Symposium
November 6 - 10, 2018, João Pessoa, Paraíba, Brazil

Verisimilar Percept Sequences Tests for Autonomous Driving Intelligent Agent Assessment

Thomio Watanabe and Denis Wolf

Mobile Robotics Laboratory
Institute of Mathematics and Computer Sciences, USP

Laboratório de Robótica Móvel
ICMC/USP - São Carlos

ICMC USP SÃO CARLOS

Safety

- Autonomous cars → improve safety
  - Human driver blamed for (94 ± 2.2)% of motor vehicle crashes in US *
  - Simple logic: remove driver → safer traffic
- What is safety and how to measure safety ?
  - Relative emotional state
  - It is not a measurable quantity !!!

* Singh, S., 2015. Critical reasons for crashes investigated in the national motor vehicle crash causation survey (No. DOT HS 812 115).

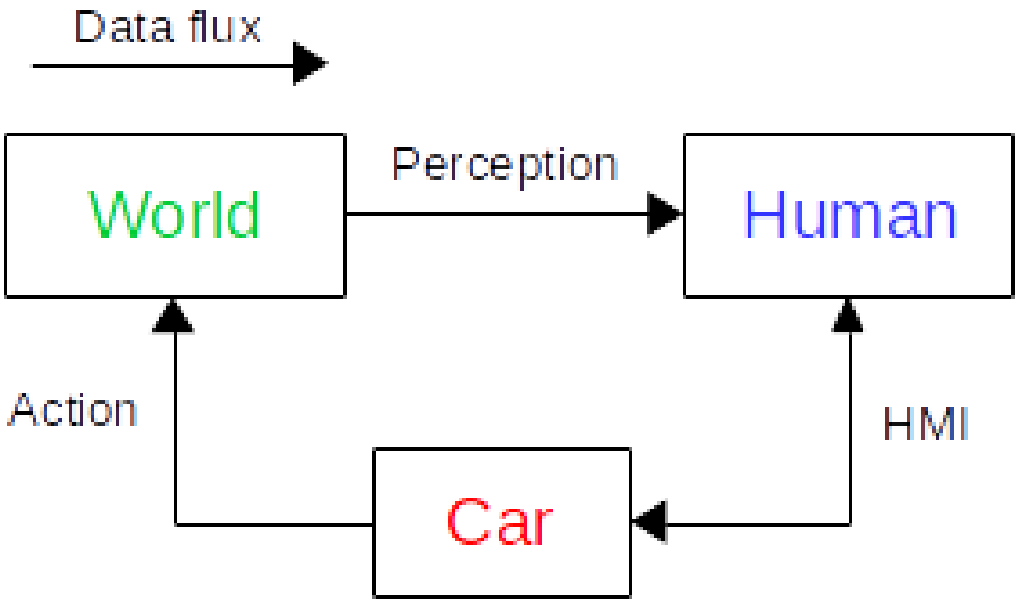
Safety

- March 23, 2018 - Uber self-driving car
  - First pedestrian kill

https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe

The Usual Car

- Human driver
  - perceive environment & controls vehicle



The Usual Car

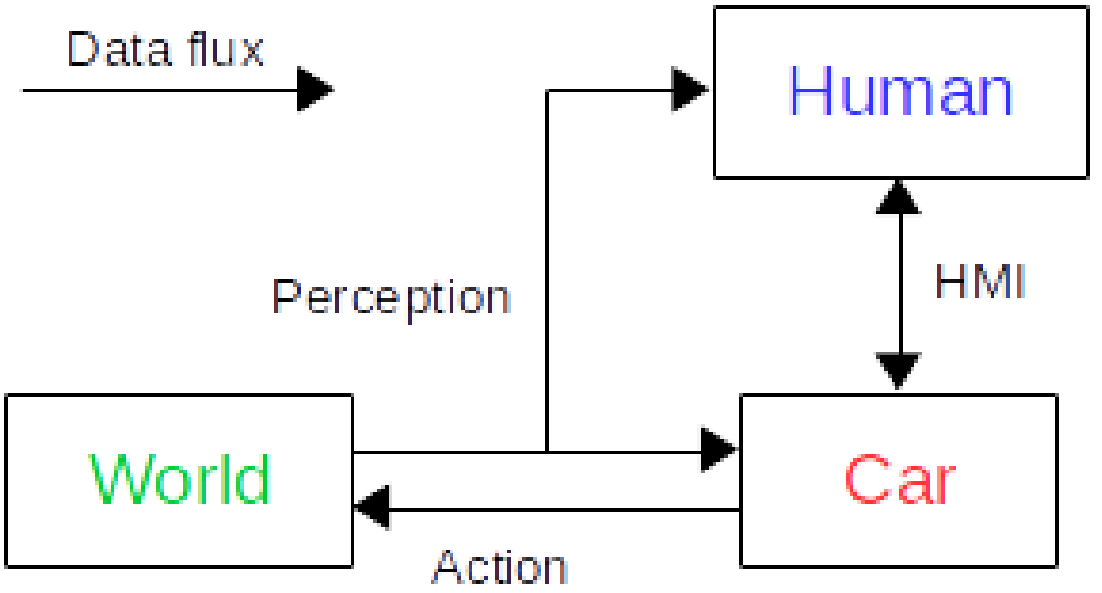
- What the human driver do ?
  - Using the eyes:
    - Identify road
    - Identify obstacles (vehicles, pedestrians, cyclists)
    - Identify traffic signs
    - objects distance
  - Arms and legs:
    - Controls steering wheel, pedals and gearbox
- Drivers must have a perfect vision and body movement.

Driver Evaluation

- Human driver evaluation:
  - Medical exams
  - Theoretical and practical tests
- Driver’s license upon approval
  - No self-driving car license

Autonomous Cars

- Human driver is optional



Car Evaluation

- Autonomous car evaluation:
  - Disengagement rate
    - Number of times of human driver intervention
    - Lower disengagement rate does not mean safer systems
  - There is no standard
    - Perception: datasets
    - System integration: field tests
    - They are not enough

Car Evaluation

- Dataset problems:
  - Cherry-picked:
    - Perfect data (images, LIDAR, radar)
    - Restricted domain representation
- Field tests problem:
  - After action review → number of accidents
    - Time-consuming and dangerous

Verisimilar Percept Sequences

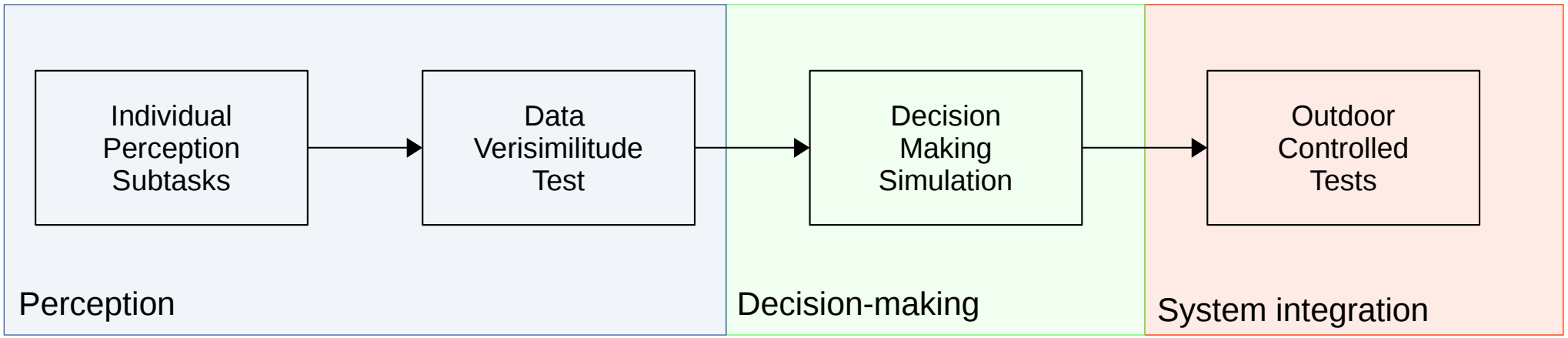
- Our proposal:
  - Intelligent Agent / Rational Agent (RA)
    - perception + decision-making
- RA as a critical system → it must not fail
  - Environment constraints evaluation
- Evaluate the perception with realistic data
  - We describe the problems that arise from realistic situations and data.

Verisimilar Percept Sequences

- Percept Sequences (PS)
  - "*the complete history of everything the agent has ever perceived*" - Russell and Norvig *
  - streams of sensors data with timestamps
  - sequential data → estimate future states
    - Partially or completely occluded objects

* S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Prentice-Hall, 2009.

Evaluation Pipeline



Evaluation Pipeline

- Individual Perception Subtasks →
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Usual dataset – evaluates performance on perception subtasks (detection / segmentation). Useful to train and compare models.

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test → Percept sequences with realistic data
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation → Percept sequences with partially and totally occluded obstacles
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence → Percept sequences with rare and unusual examples
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption → Percept sequences with incomplete data
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption → Percept sequences with noisy data
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks → Percept sequences with simulated data attacks (adversarial examples)
- Decision-making Simulation
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation → Percept sequences representing task environments
- Outdoor Controlled Tests

Evaluation Pipeline

- Individual Perception Subtasks
- Data Verisimilitude Test
  - Data estimation
  - Data absence
  - Data flux interruption
  - Data corruption
  - Data attacks
- Decision-making Simulation
- Outdoor Controlled Tests → Mock-up field tests to validate system integration.

Conclusion

- Stimulate discussion about the topic
  - Model and list of rational agents problems
  - Autonomous car evaluation → refined over time
- Future work:
  - Percept sequences examples
  - Each step can be created separately

Thanks !

thomio.watanabe@usp.br