Statistical Model Checking Applied on Perception and Decision-making Systems for Autonomous Driving

J. Quilbeuf [1] M. Barbier [2,3] L. Rummelhard [3]
C. Laugier [2] A. Legay [1] T. Genevois [2]
J. Ibañez-Guzmán [3] O. Simonin [4]

[1]Univ Rennes, Inria, F-35000, Rennes, France.

[2]Univ. Grenoble Alpes, Inria, Chroma, F38000 Grenoble, France.

[3]Renault S.A.S, 1 av. du Golf, 78288 Guyancourt, France.

[4]INSA Lyon, CITI Lab., 6 avenue des Arts, 69680 Villeurbanne, France.

September 30, 2018

Introduction

Introduction

Rationale

Classical approaches for validation in the automotive industry:

- Vehicle-in-the-loop platform to test interactions between a human and the system in dangerous situations [Bokc 2007]
- Hardware-in-the-loop to test interactions between an embedded system and the dynamics of the vehicle [Hwang 2006]

Rationale

Classical approaches for validation in the automotive industry:

- Vehicle-in-the-loop platform to test interactions between a human and the system in dangerous situations [Bokc 2007]
- Hardware-in-the-loop to test interactions between an embedded system and the dynamics of the vehicle [Hwang 2006]

Not enough for autonomous vehicle systems that target SAE level 3 and higher:

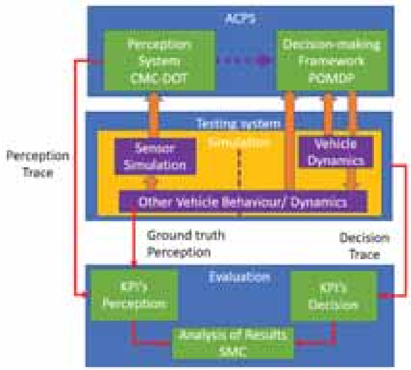
- No driver
- Interactions between systems
- Uses learning and probabilities
- Many scenarios

Problematic

- In the context of autonomous vehicles, what process can be applied to validate a system that enable high level of autonomy?
- How to formulate requirements for validation?
- What are the simulation tools requirements for validation?

Contributions

- Application of statistical model checking on two key elements of autonomous vehicle systems:
  - Decision-making
  - Perception
- Key performances indicators (KPI) for systems or scenarios
- Analysis of SMC results (i.e Probability of meeting a KPI)



Statistical Model Checking

Overview

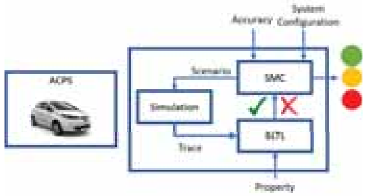
It provides an intermediate between test and exhaustive verification by relying on statistics [Sen 2005]

Goal

Evaluation of the probability to meet a property (or Key Performance Indicators) out of many executions

SMC needs:

- Stochastic simulations
- Stochastic models
- Scenario variations



Principle

Monte-Carlo formulation

$$\hat{p} = \frac{1}{N}\sum_{1}^{N} f(ex_i) \quad where \quad f(ex_i) = \begin{cases} 1 & \text{if} \quad ex_i \models \phi \\ 0 & \text{otherwise} \end{cases}$$

$\hat{p}$ estimation of the probability
N number of simulations

Chernoff bound

$$Pr(|p - \hat{p}| \leq \epsilon) \geq 1 - \delta$$

$$N > \frac{log(\frac{2}{\delta})}{2\epsilon^2}$$

p the probability to evaluate
The estimation error is bounded by ϵ the error with a probability $1 - \delta$

Bounded Linear Temporal Logic

Formula to express if a property ϕ is found within an execution trace that is a sequence of state p with a stamp t

Syntax [Zuliani 2013]

	logical		temporal	
$\phi \; ::= p$	$\phi \vee \phi$	$\neg\phi$	$\phi \, U_{\leq t} \, \phi$	$X_{\leq t} \, \phi$
predicate	disjunction	negation	Until	Next

Example

$$F_{\leq d} crossed$$

Finally before d time elapsed crossed is always false

A first validation application: CMCDOT perception system

Conditional Montecarlo Dense Occupancy Tracker

- Estimate Spatial occupancy for each cell of the grid $P(O|Z)$ (Static, Dynamic, Empty, Unknown)
- Grid update is performed in each cell in parallel (using BOF equations)
- Reason at the Grid level (i.e. no object segmentation at this reasoning level)
- Dense Occupancy Tracker (Object level, Using particles propagation and ID)[Rummelhard 2015]

Click!

Time-To-Collision computation

Click!

Simulation

Features

- Precise volume, shape, surface
- Atmospheric condition
- Ground truth as occupancy grid

Tools

- ROS: Robotic middleware
- Gazebo

Click!

KPI CMCDOT

System driven KPI

Problem: The nature of the output of the CMCDOT is a probabilistic grid what is the ground truth for that
Solution: Observe the result of an application of the CMCDOT

KPI CMCDOT

System driven KPI

Problem: The nature of the output of the CMCDOT is a probabilistic grid what is the ground truth for that
Solution: Observe the result of an application of the CMCDOT

TTC KPI

$G_{\leq t}(\texttt{real\_coll}_i \Rightarrow (1 - \texttt{cmcdot\_risk}) < \tau) \wedge (\neg \texttt{real\_coll}_i \Rightarrow \texttt{cmcdot\_risk}) < \tau)$
This property states that if there is a risk of collision, the probability returned by CMCDOT must be high enough.
Conversely, if there is no risk of collision, the probability returned by CMCDOT must be small enough.

Results

Work in progress
Difficulties:

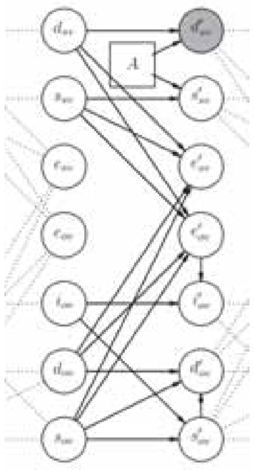
- Generate a ground truth for occupancy grids
- determinism problem with ROS
- No simulators with all the requirement available

A second validation application: a decision-making system

POMDP applied on road intersection crossing

Partially Observable Markov Decision Process

- Consider uncertainties
- Reward function uses:
  - Variation from reference speed
  - Risk
  - Acceleration changes
- Actions are a range of accelerations and decelerations
- Online solver for real time but partial policy value estimation



Simulation requirements

To test a decision-making system, the simulation must feature:

Interactive behaviour

Vehicles within the simulation environment must react to actions chosen by the ego vehicle

Scenario variations

As many parameters as required to reproduce real life scenes must be configurable (e.g vehicle speeds, traffic signs)

Uncertainties

Observations returned by the simulation must reproduce errors and uncertainties from perception system and vehicle dynamics

Simulation requirements

To test a decision-making system, the simulation must feature:

Interactive behaviour

Vehicles within the simulation environment must react to actions chosen by the ego vehicle

Scenario variations

As many parameters as required to reproduce real life scenes must be configurable (e.g vehicle speeds, traffic signs)

Uncertainties

Observations returned by the simulation must reproduce errors and uncertainties from perception system and vehicle dynamics

Solution retained: Scaner (automotive grade simulators)

Decision execution

KPI for decision-making for crosscutting scenarios

Scenario driven approach

Metrics are defined from highway code or from what can be observed of the situation

Name	Description	Unit
t	Timestamp or time elapsed	s
nc_stops	Number of stops in the non-critical area	
c_stops	Number of stops in the critical area	
t_nc_stops	Duration of stops in non-critical area	s
t_c_stops	Duration of stops in critical area	s
acc	Acceleration	ms^{-2}
crossed	True if intersection is crossed	

KPI for decision-making for crosscutting scenarios

Scenario driven approach

Metrics are defined from highway code or from what can be observed of the situation

Name	Description	Unit
t	Timestamp or time elapsed	s
nc_stops	Number of stops in the non-critical area	
c_stops	Number of stops in the critical area	
t_nc_stops	Duration of stops in non-critical area	s
t_c_stops	Duration of stops in critical area	s
acc	Acceleration	ms^{-2}
crossed	True if intersection is crossed	

$$N > \frac{log(\frac{2}{\delta})}{2\epsilon^2}$$

With $N = 800$ and $\delta = 0.01$ we have $\epsilon = 0.0137$

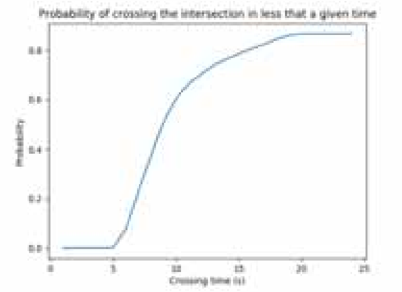
KPI: Crossing time

BLTL Statement

$$F_{\leq d} \textit{crossed}$$

The vehicle crossed the intersection within the bound d

- The intersection is never crossed in $5s$ or less
- Most likely the intersection is crossed in $10s$
- There is a probability of 0.1 that the vehicle does not cross



KPI Stopping in critical area

BLTL Statement

$$F_{\leq d} t\_c\_stops$$
$$F_{\leq d} t\_nc\_stops$$

- Unlikely to stop in the critical area
- Stopping before the intersection has a probability of occurring of 0.25
- The decision making system is able to slow down to let the other vehicle cross

Conclusion

Validation

Statistical model checking offers information on the system as well as how confident measures are

Simulation

Even if many simulators exist, features required for validation are not often present.

Requirement specification

Key Performance Indicator formulate as bounded linear temporal logic creates a rich syntax for validation requirement

Further works

- Combine the analysis of the decision-making and perception to understand their effect on each other.
- Create KPI that depend on the state of other vehicle.

References

- T. Bokc and M. Maurer and G. Farber (2007)
  Validation of the Vehicle in the Loop (VIL); A milestone for the simulation of driver assistance systems
  2007 IEEE Intelligent Vehicles Symposium
- T. Hwang and J. Roh and K. Park and J. Hwang and K. H. Lee and K. Lee and S. j. Lee and Y. j. Kim (2006)
  Development of HILS Systems for Active Brake Control Systems
  2006 SICE-ICASE International Joint Conference
- Koushik Sen and Mahesh Viswanathan and Gul Agha (2005)
  On statistical model checking of stochastic systems
  Proceedings of the 17th International Conference on Computer Aided Verification
- Zuliani, Paolo and Platzer, André and Clarke, Edmund M." (2013)
  Bayesian statistical model checking with application to Stateflow/Simulink verification
  Formal Methods in System Design oct 2013
- Rummelhard, Lukas and Negre, Amaury and Laugier, Christian" (2015)
  Conditional Monte Carlo Dense Occupancy Tracker
  18th IEEE International Conference on Intelligent Transportation Systems

The End

Use case: Road intersection crossing

Accidentologie

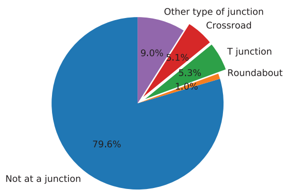


Figure: 20 % of accidents at junction

Figure: Google car after accident at a road intersection

Use case: Road intersection crossing

Accidentologie

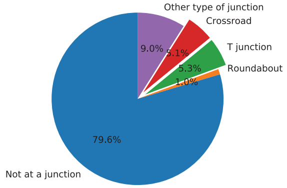


Figure: 20 % of accidents at junction

Figure: Google car after accident at a road intersection

Difficulties for decision and perception

- Uncertainties
- Driver's behaviour

KPI: bounded acceleration

BLTL Statement

$$G_{\leq t} F_{\leq 1} Acc \leq b.$$

Acc will be smaller than *b* in less than 1s. In other words, it is not possible that $Acc > b$ for more than 1s. The value of the bound *b* is defined w.r.t. the metric considered.

- An acceleration of $2m/s^2$ is highly likely to happen at least once
- The probability that the acceleration is below $2m/s^2$ is 0.6
- The system has two acceleration spikes for short-time periods