Preemptive Detection of Unsafe Motion Liable for Hazard

Masataka Nishi
Hitachi Research Laboratory, Hitachi Ltd.

 <masataka.nishi.en@hitachi.com>

High-level concerns

What is safety?

- How infrequently a bad consequence could occur.

Developer's concern

- How to support a safety claim <1 crash/1mil.[km] in a quantifiable way?

Regulator's concern

- Is the vehicle formally verifiable, certifiable and incrementally improved?
- How to find a technical defect at pre-market stage or after accumulation of accidents?

Insurer's concern

- How to determine who is liable? Based on what legal scheme? What is legally valid evidence?

Table 1. Examples of Miles and Years Needed to Demonstrate Autonomous Vehicle Reliability

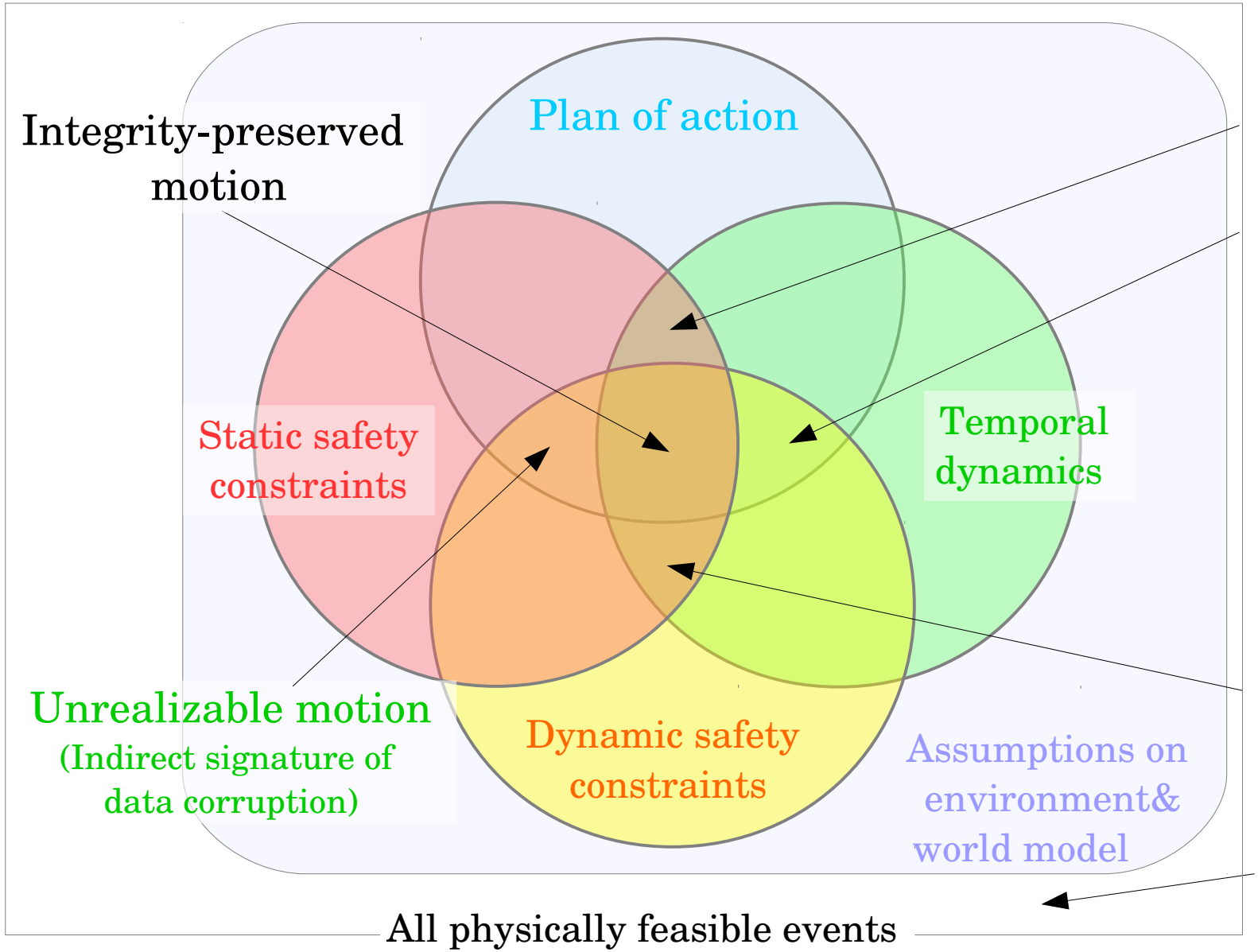
		Benchmark Failure Rate		
Statistical Question	How many miles (years[a]) would autonomous vehicles have to be driven...	(A) 1.09 fatalities per 100 million miles?	(B) 77 reported injuries per 100 million miles?	(C) 190 reported crashes per 100 million miles?
	(1) without failure to demonstrate with 95% confidence that their failure rate is at most...	275 million miles (12.5 years)	3.9 million miles (2 months)	1.6 million miles (1 month)
	(2) to demonstrate with 95% confidence their failure rate to within 20% of the true rate of...	8.8 billion miles (400 years)	125 million miles (5.7 years)	51 million miles (2.3 years)
	(3) to demonstrate with 95% confidence and 80% power that their failure rate is 20% better than the human driver failure rate of...	11 billion miles (500 years)	161 million miles (7.3 years)	65 million miles (3 years)

N. Kalra, S. M. Paddock, "How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?".RR1478, RAND Corporation

 <masataka.nishi.en@hitachi.com>

Built-in defects: Loss of functional integrity

Direction 1: Detect loss of functional integrity by checking SATisfiability of these at run-time



 <masataka.nishi.en@hitachi.com>

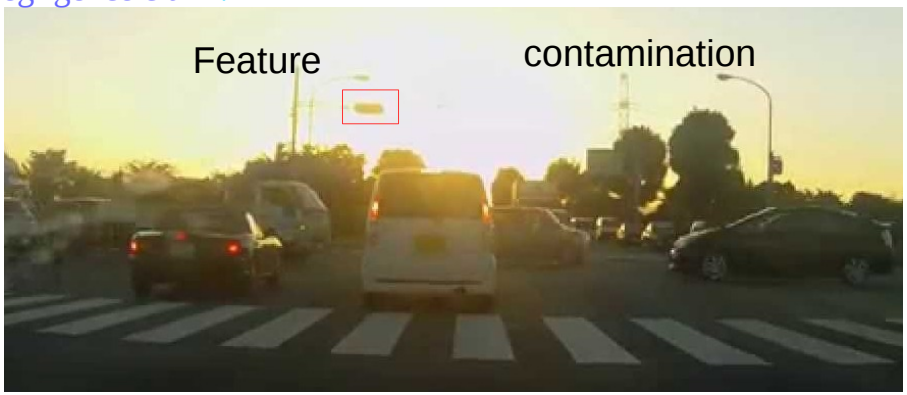
Externality of the risk of hazard

Direction 2: Determine liability based on a root-cause of evidently bad consequence

Plaintiff/Driver&Auto insurer:
"Unintended acceleration by itself!"
Defendant/OEM: "No evidence on record. You're responsible for taking control as written in the disclaimer."

Plaintiff/Injured party: "My house got damaged!"
Defendant/Driver&Insurer: "Not my fault. The vehicle did it!"

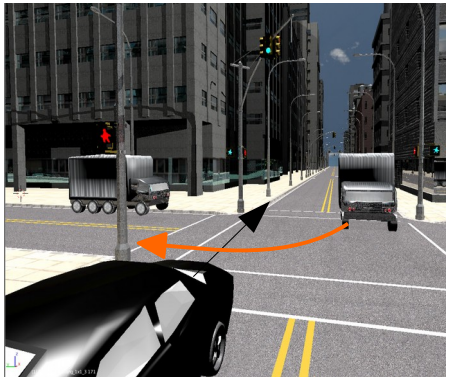
Plaintiff/Regulator: "Recall all! Technical defect is repeatable!"
Defendant/OEM: "Reached the technological limit. Exempt from negligence claim."



Defect in the vehicle?

Root-cause built in the environment?

Plaintiff/Driver: "Go or brake?"
Defendant/Truck: "I turned, as you flashed."



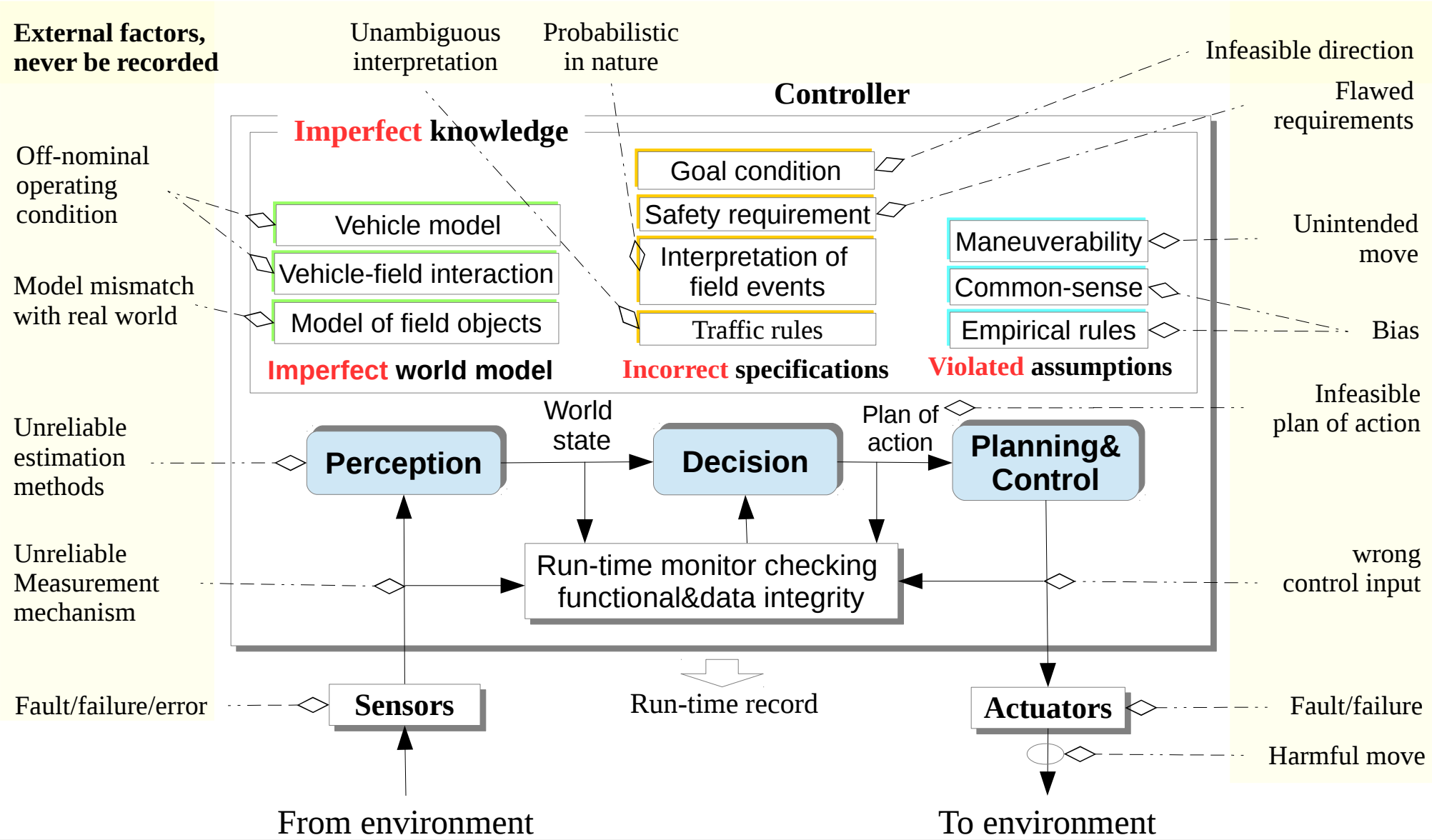
Plaintiff/OEM: "Repair it!"
Defendant/Roadway Service: "Machine vision should keep attention. Just avoid it."

Plaintiff/Driver,OEM&auto insurer:
"Lane departure due to poor lane marking"
Defendant/Roadway Service: "We can guess effortlessly."

 <masataka.nishi.en@hitachi.com>

Single-points of failure: what if external factors=root-cause?

Direction3: Understand fundamental limits of determining liability based on root-cause analysis.
Formal root-cause analysis by solving MAXSAT is difficult, if only partially observable.



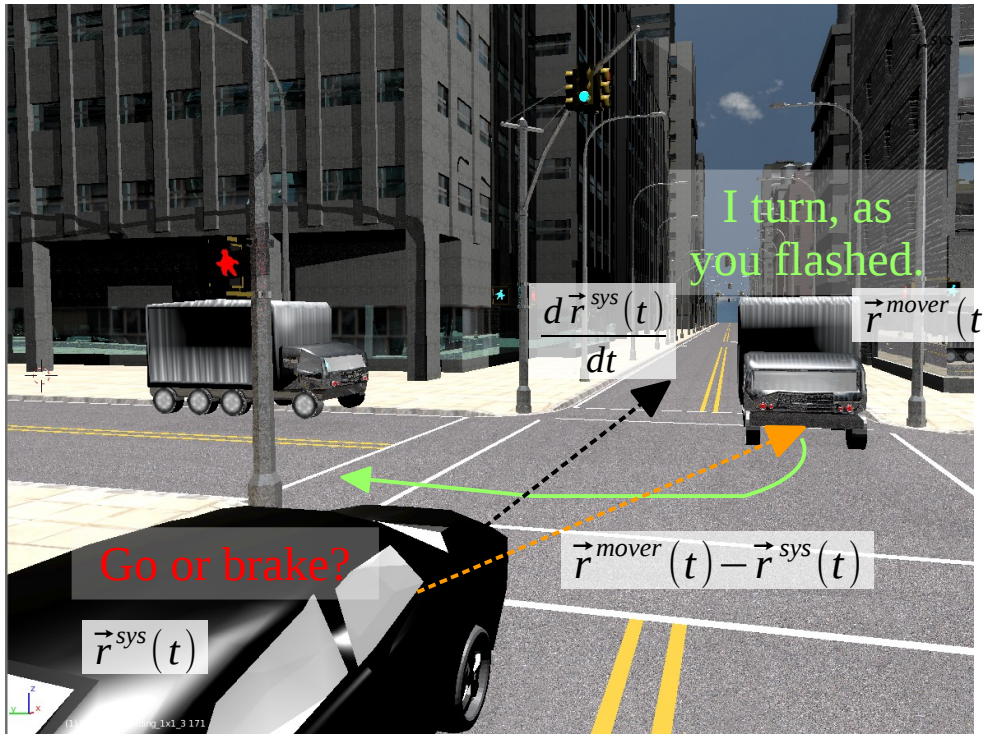
 <masataka.nishi.en@hitachi.com>

Sidestepping the fundamental limits on determining liability

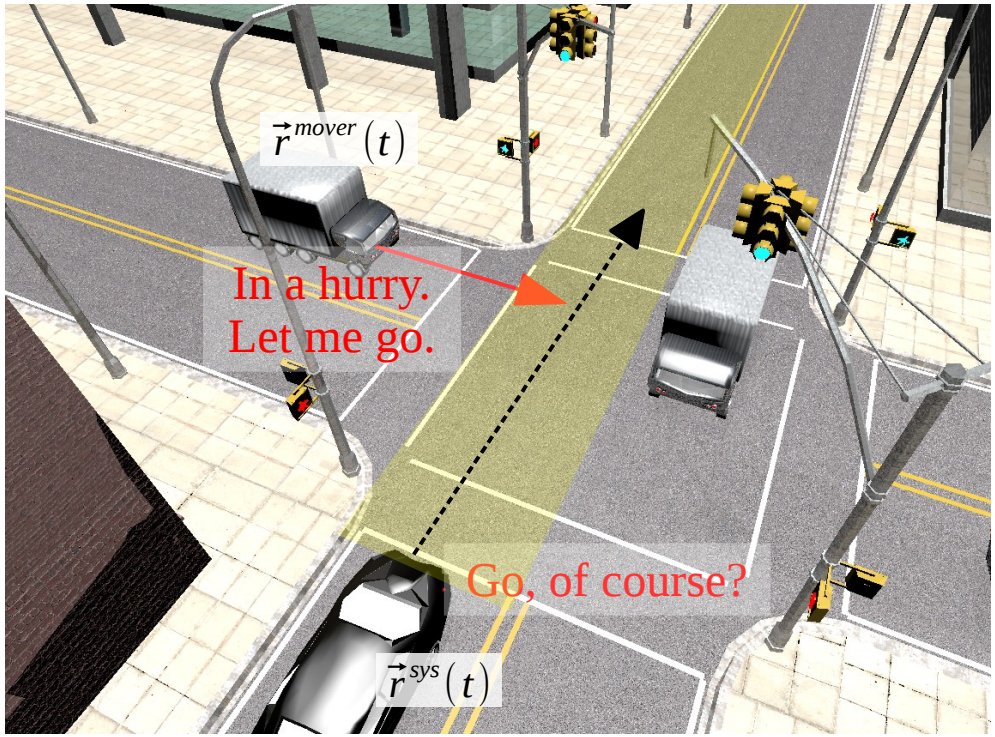
Direction 4: Determine liability using only observable states, based on contribution to hazard

$$haz(r^{sys}(t), r^{mover}(t)) \equiv \underbrace{\left|r^{mover}(t) - r^{sys}(t)\right|^2 \le \epsilon^2}_{\text{Premise of collision}} \wedge \{ \underbrace{\frac{d\,r^{sys}(t)}{dt} \frac{r^{mover}(t) - r^{sys}(t)}{\left|r^{mover}(t) - r^{sys}(t)\right|} \ge \delta}_{\text{Contribution to hazard}} \vee \neg \text{right-of-way} \}$$

If you hit without evasive move,
you are at least partially liable for the crash.



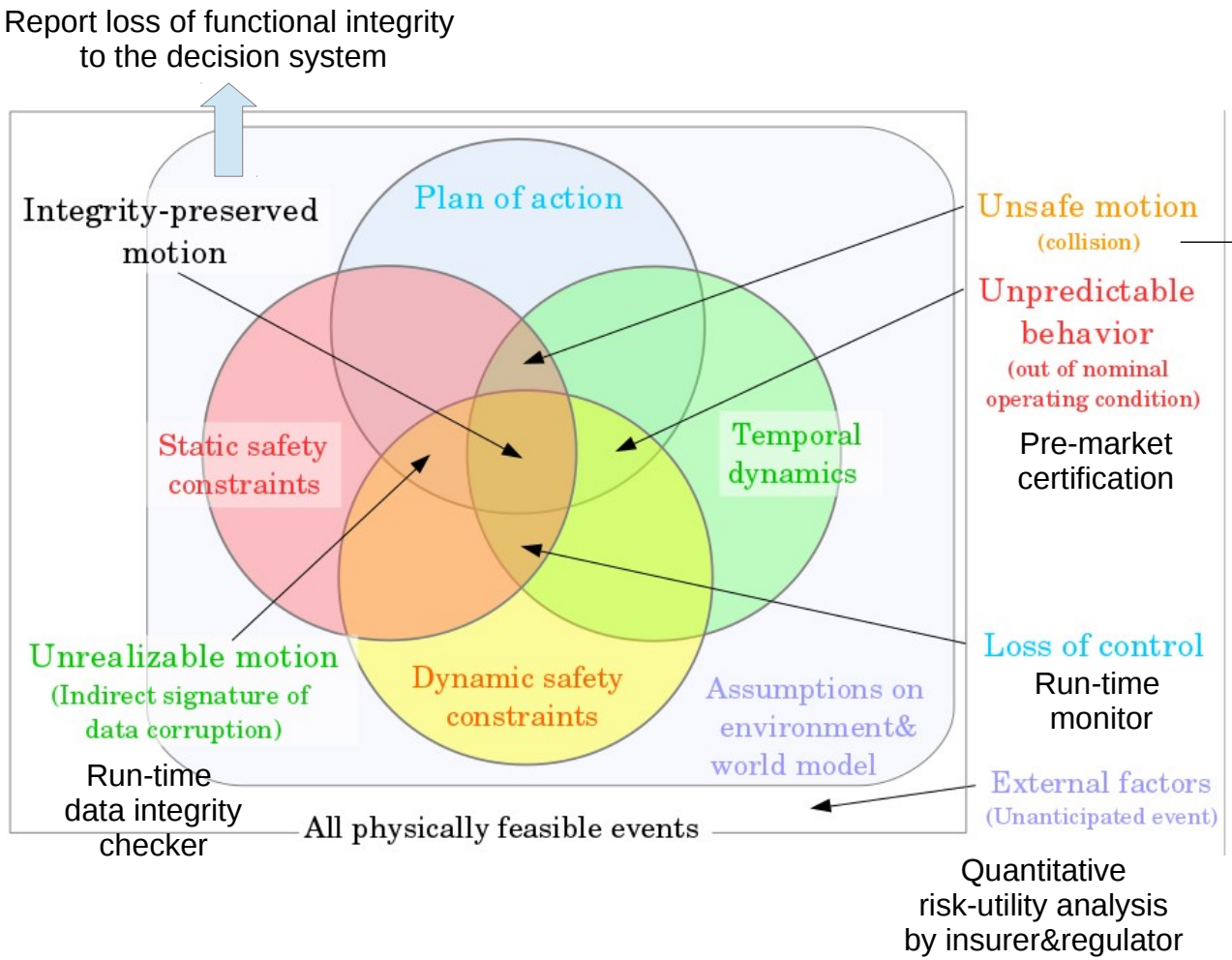
The truck doesn't have a right-of-way,
the truck is totally liable for crash.



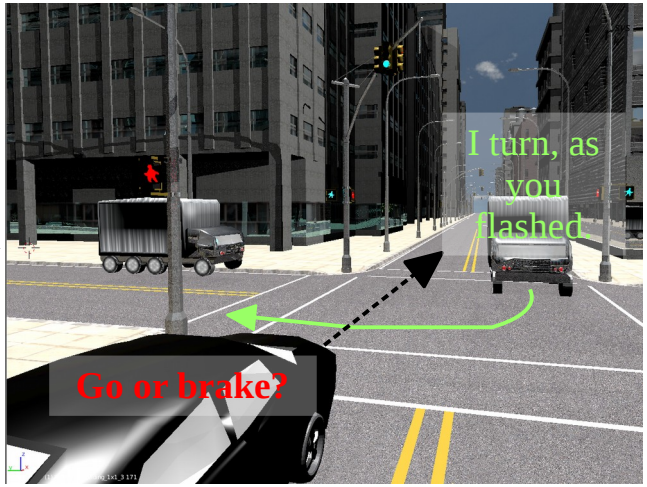
 <masataka.nishi.en@hitachi.com>

What root-cause to detect preemptively? Non-stochastic?

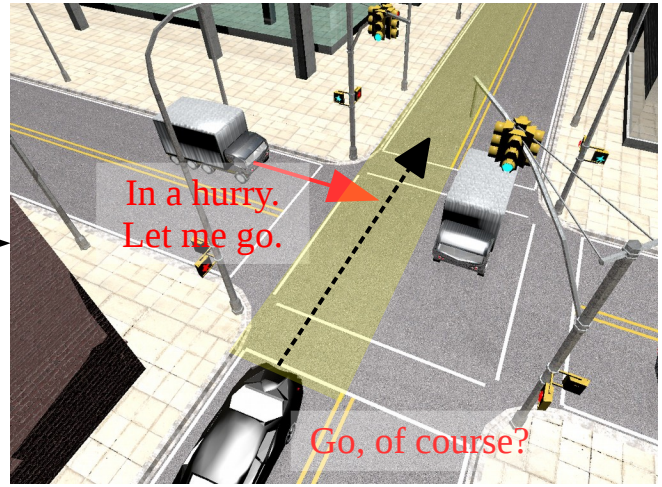
Crash frequency becomes quantifiable, by preemptively detecting non-stochastic root-causes.



If you hit without evasive move...



or you don't have a right-of-way,



Preemptive detection of infeasible plan of action/Loss-of-Control

World state

$\boldsymbol{X}_T \equiv \{\boldsymbol{x}_{0,\ldots}, \boldsymbol{x}_T\}, \boldsymbol{U}_T \equiv \{\boldsymbol{x}_{0,\ldots}, \boldsymbol{x}_T\}$

$0 \le k < T, \boldsymbol{x}_k :\equiv \{\vec{r}_k, v_k, \theta_k\} \in \boldsymbol{R}^4, \boldsymbol{u}_k \equiv \{a_k, \tau_k\} \in \boldsymbol{R}^2$

{position, velocity, yaw angle} {acceleration, yaw rate}

Temporal dynamics

$$\boldsymbol{F}^{sys}(\boldsymbol{X}_T, \boldsymbol{U}_T) \equiv \begin{cases} -\vec{r}_{k+1} + \vec{r}_k + v_{k+1}\begin{bmatrix}\cos(\theta_{k+1})\\ \sin(\theta_{k+1})\end{bmatrix}\Delta t = \vec{0}, 0 \le \forall k \le T \\ -\begin{bmatrix}v_{k+1}\\ \theta_{k+1}\end{bmatrix} + \begin{bmatrix}v_k\\ \theta_k\end{bmatrix} + \begin{bmatrix}a_{k+1}\\ \tau_{k+1}\end{bmatrix}\Delta t = \vec{0}, 0 \le \forall k \le T \end{cases}$$

Static safety constraint

$$\boldsymbol{S}^{sys}(\boldsymbol{X}_T, \boldsymbol{U}_T) \equiv \begin{bmatrix}\underline{\boldsymbol{u}}\\ \underline{v}\end{bmatrix} \le \begin{bmatrix}\boldsymbol{u}_k\\ v_k\end{bmatrix} \le \begin{bmatrix}\overline{\boldsymbol{u}}\\ \overline{v}\end{bmatrix} \wedge \underline{\boldsymbol{du}} \le \frac{1}{\Delta t}(\boldsymbol{u}_{k+1} - \boldsymbol{u}_k) \le \overline{\boldsymbol{du}}$$

Mechanical limit of steering/acceleration

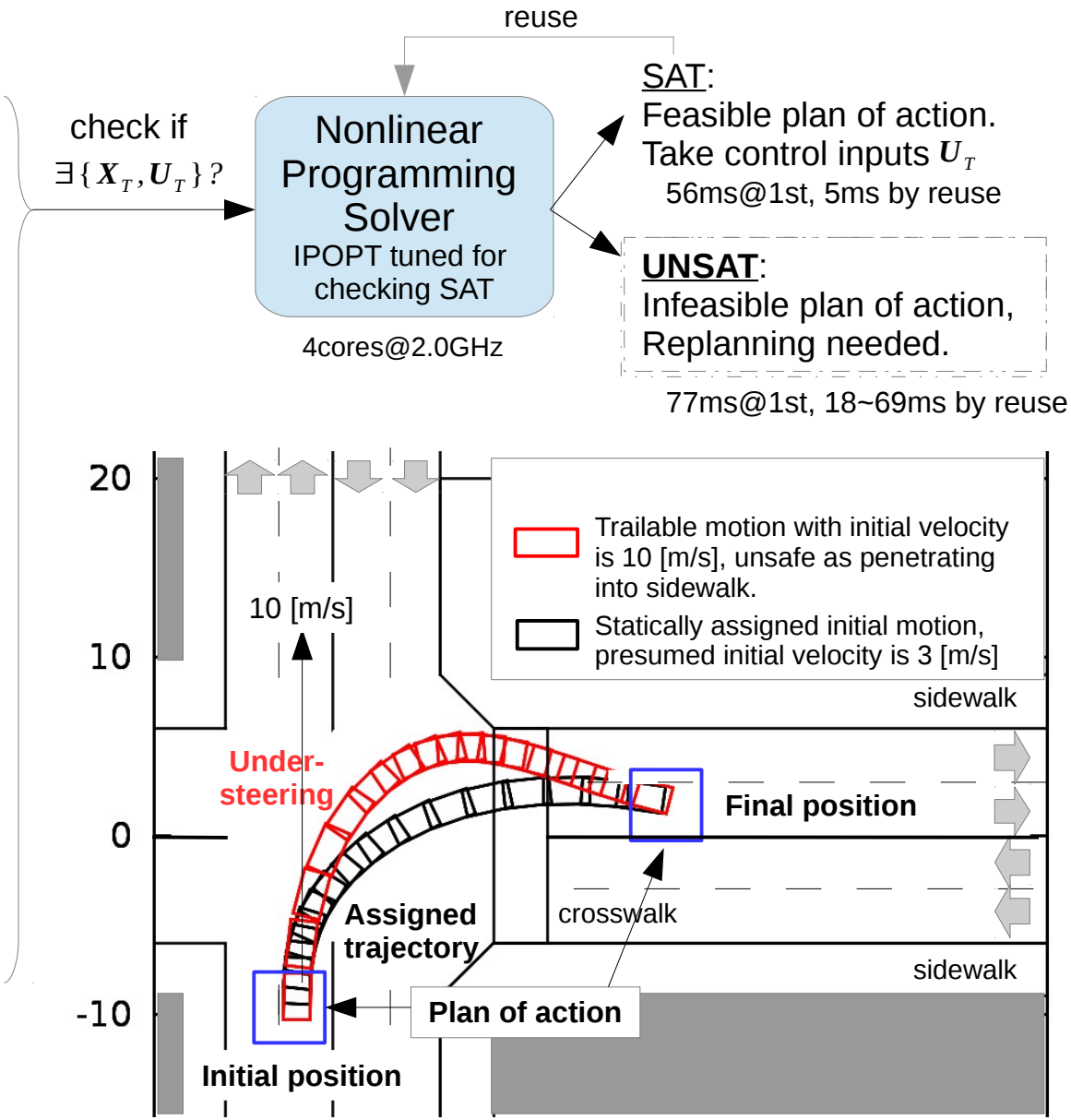
Dynamic safety constraint

$\boldsymbol{D}(\boldsymbol{X}_T) \equiv \{\vec{r}_k \in R_{\text{RIGHT-OF-WAY}} \mid 0 \le \forall k \le T\}$

Plan of action, route navigator demands to turn

$\vec{r}_0 \in R_{INIT}, \vec{r}_T \in R_{GOAL}, \{\vec{r}_k \in R_{\text{REFERENCE\_PATH}} \mid 0 \le \forall k \le T\}$

Terminal conditions Simple path-following



Preemptive detection of unsafe motion

by checking satisfiability of an adversarial motion planning from the mover's standpoint

Plan of action

$$\vec{r}_0^{\,sys} \in R_{INIT}, \vec{r}_T^{\,sys} \in R_{GOAL}, \{\, \vec{r}_k^{\,sys} \in R_{\mathrm{REFERENCE\_PATH}} \mid 0 \le \forall k \le T \,\}$$

World state

$$\boldsymbol{X}_{i,T}^{mover} \equiv \{\, \boldsymbol{x}_{i,0}^{mover}, \ldots, \boldsymbol{x}_{i,T}^{mover} \,\}, \boldsymbol{U}_{i,T}^{mover} \equiv \{\, \boldsymbol{x}_{i,0}^{mover}, \ldots, \boldsymbol{x}_{i,T}^{mover} \,\}$$

$$\boldsymbol{x}_{i,k}^{mover} :\equiv \{\, \vec{r}_{i,k}, v_{i,k}, \theta_{i,k} \,\} \in \boldsymbol{R}^4, \boldsymbol{u}_{i,k}^{mover} \equiv \{\, a_{i,k}, \tau_{i,k} \,\} \in \boldsymbol{R}^2$$

{position, velocity, yaw angle} {acceleration, yaw rate}

Temporal dynamics

$$\boldsymbol{F}^{mover}\left(\boldsymbol{X}_{i,T}^{mover}, \boldsymbol{U}_{i,T}^{mover}\right) \equiv \begin{cases} -\vec{r}_{k+1} + \vec{r}_{i,k} + v_{i,k+1} \begin{bmatrix} \cos(\theta_{i,k+1}) \\ \sin(\theta_{i,k+1}) \end{bmatrix} \Delta t = \vec{0} \\ -\begin{bmatrix} v_{i,k+1} \\ \theta_{i,k+1} \end{bmatrix} + \begin{bmatrix} v_{i,k} \\ \theta_{i,k} \end{bmatrix} + \begin{bmatrix} a_{i,k+1} \\ \tau_{i,k+1} \end{bmatrix} \Delta t = \vec{0} \end{cases}$$

Assumption on maneuverability of mover

$$\boldsymbol{S}^{mover}\left(\boldsymbol{X}_{i,T}^{mover}, \boldsymbol{U}_{i,T}^{mover}\right) \equiv \underline{\begin{bmatrix} \underline{u_i} \\ \underline{v_i} \end{bmatrix} \le \begin{bmatrix} u_{i,k} \\ v_{i,k} \end{bmatrix} \le \begin{bmatrix} \overline{u_i} \\ \overline{v_i} \end{bmatrix}} \wedge$$

Physical limit of rotation/acceleration

$$\underline{\boldsymbol{du}_i} \le \frac{1}{\Delta t}\left(\boldsymbol{u}_{i,k+1} - \boldsymbol{u}_{i,k}\right) \le \overline{\boldsymbol{du}_i}$$

Smoothness of the mover's motion

Dynamic safety constraint

$$\boldsymbol{D}\left(\boldsymbol{X}_T^{sys}, \boldsymbol{X}_{i,T}^{mover}\right) \equiv \neg \bigvee_{0 \le k \le T} haz\left(\vec{r}_k^{\,sys}, \vec{r}_{i,k}\right)$$

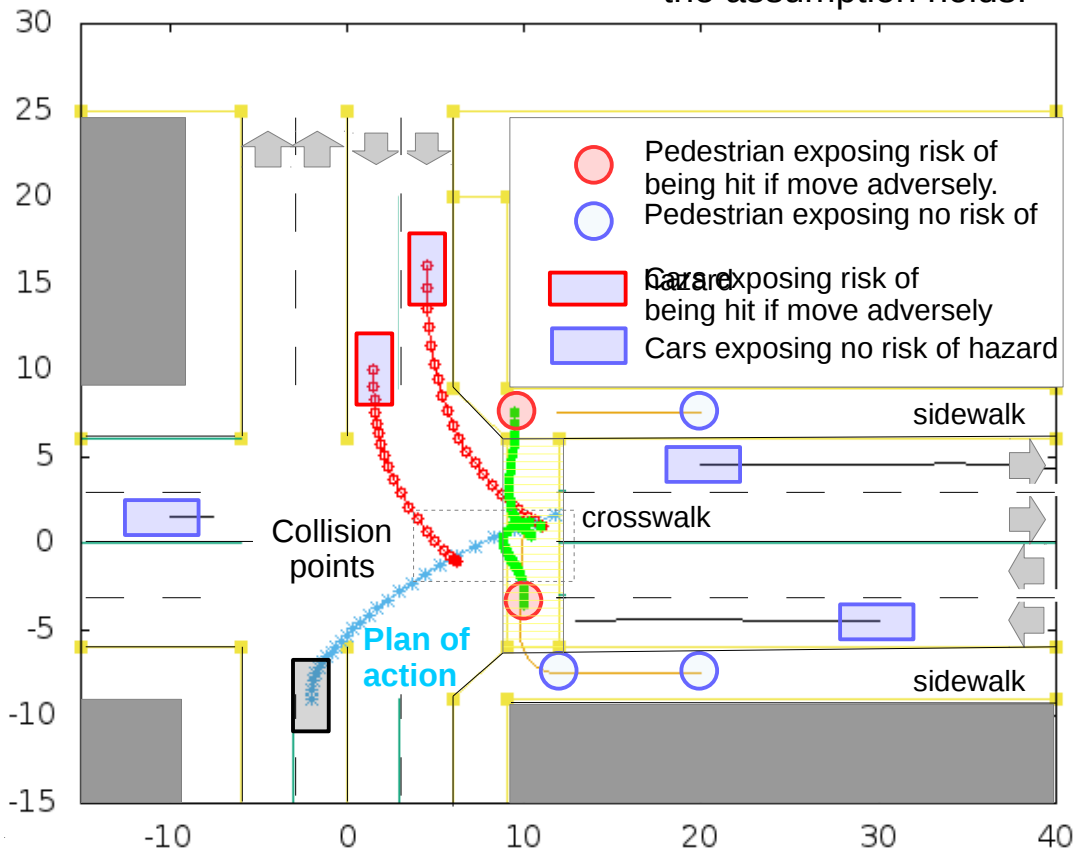
check if $\exists \{\, \boldsymbol{X}_{i,T}^{mover}, \boldsymbol{U}_{i,T}^{mover} \,\}$?

Nonlinear Programming Solver

SAT:
Feasible adversarial move.
Risk of hazard detected

10 movers at once in parallel
71~89ms@1st, 28-59ms by reuse

UNSAT:
No risk of hazard **UNTIL** the assumption holds.



Determining liability in a legal context

Offense strategies

Payment for victims and owners of defective product

(Strict) liability : Unconditional compensation for damage from defects and failure of warning

Negligence : Protection of vehicle owner against conditioned design fault

- Foresee at least known risk of hazard
- Act responsibly to avoid the hazard

Misrepresentation : False/misleading information on functionality & the risk of hazard

Breach of warranty : Failure of providing the stated functionality when used in foreseeable ways.

Defense strategies

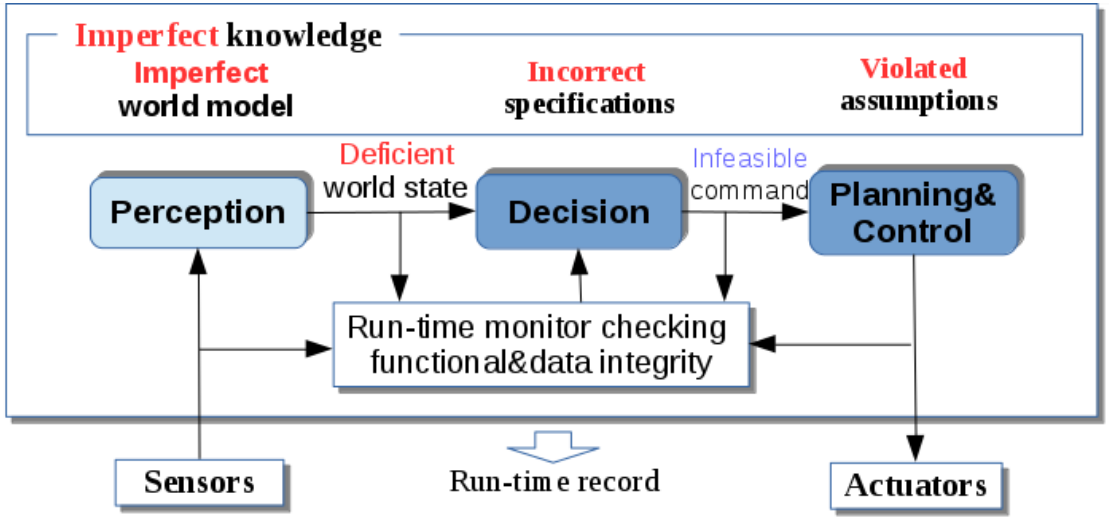
For waiver/reduction of OEM's liability

Contributory negligence : Insured party is partially liable, if contributes to the hazard.

Consent/assumption of risk : Signed acknowledgment of the known&expressed risk of hazard

Necessity, limitation : Exemption, if explicitly stated conditions hold.

: Force Majeure clause, if externality/unforeseeable/irresistibility holds.

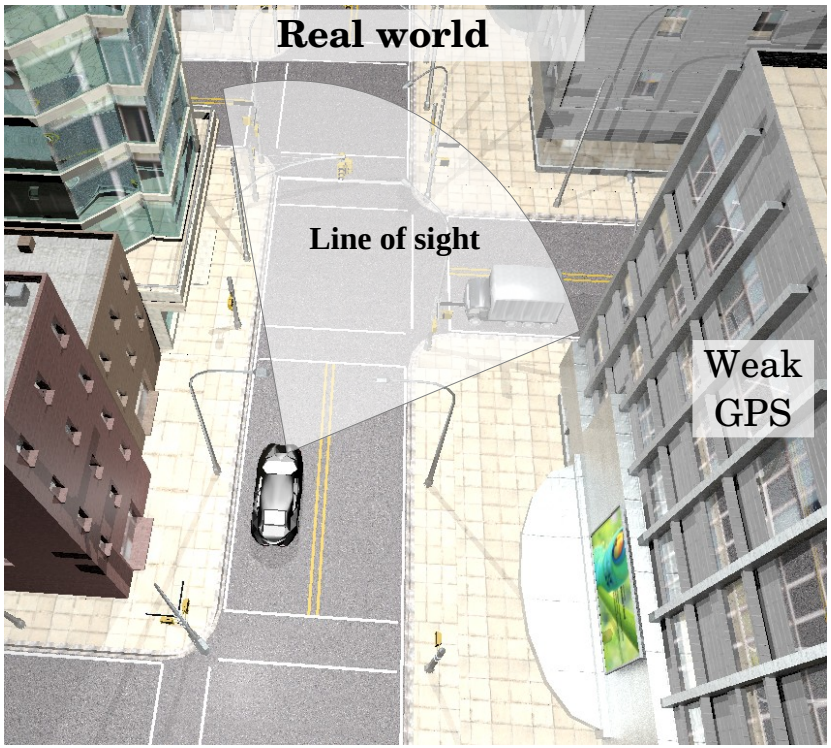


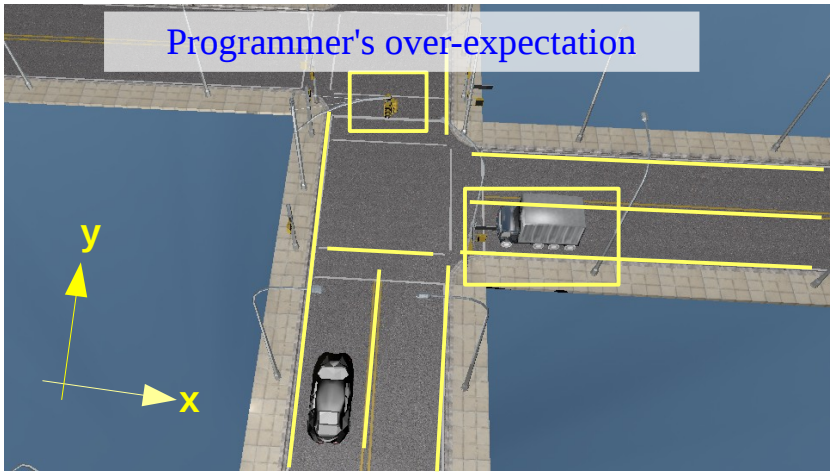
Technical summary and implications

- Safety = How infrequently a bad consequence could occur.
- Liability = Each stakeholder's responsibility for predicting and avoiding the bad consequence.
- Proposals for determining liability and reducing crash frequency over time by design
  - Direction 1: Detect loss of functional integrity
    - Certification examiner must verify if the detection&recording function works.
  - Direction 2: Determine liability based on a root-cause of evidently bad consequence
  - Direction 3: Understand fundamental limits of determining liability based on the root-cause analysis.
    - Insurer should serve as a semi-independent auditor responsible for identifying a root-cause.
  - Direction 4: Determine liability using only observable states, based on contribution to hazard
    - If you hit without evasive move or without right-of-way, then you are at least partially liable.
- Detection mechanism
  - Detect loss of functional integrity event
  - Preemptive detection of infeasible plan of action, avoiding loss of control
  - Preemptive detection of unsafe motion, avoiding getting dynamic safety constraints violated.
    - Numerically check satisfiability using Nonlinear Programming solver

 <masataka.nishi.en@hitachi.com>

Limited observability of world state





World state consists of

Directly measurable part: Y

Sensors produces, but sometimes get unavailable/unreliable.

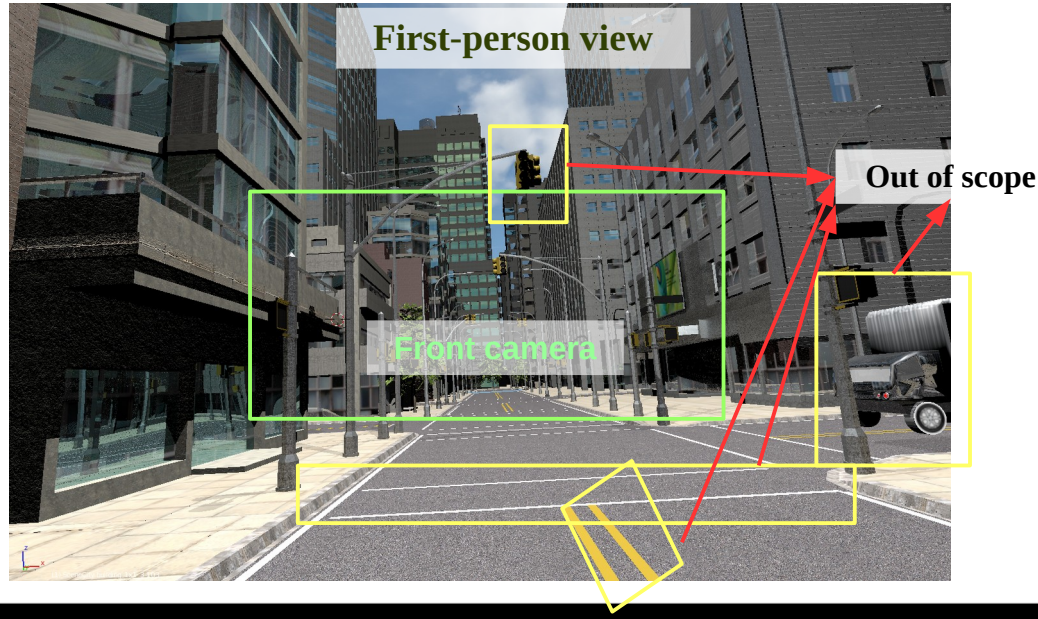
Unmeasurable part : X

We can estimate it using a world model, if observable.

Time in 100 [ms]	0	1	2	..	50
World state: Y[0]					error
Y[...]	Y: Directly measurable part				
Y[10]			unavailable		
X[0]					
X[...]	X: Unmeasurable part				
X[100]	unknown				
Z[0.100]	Not built in the world model				

Observable part

Unobservable part



 <masataka.nishi.en@hitachi.com>