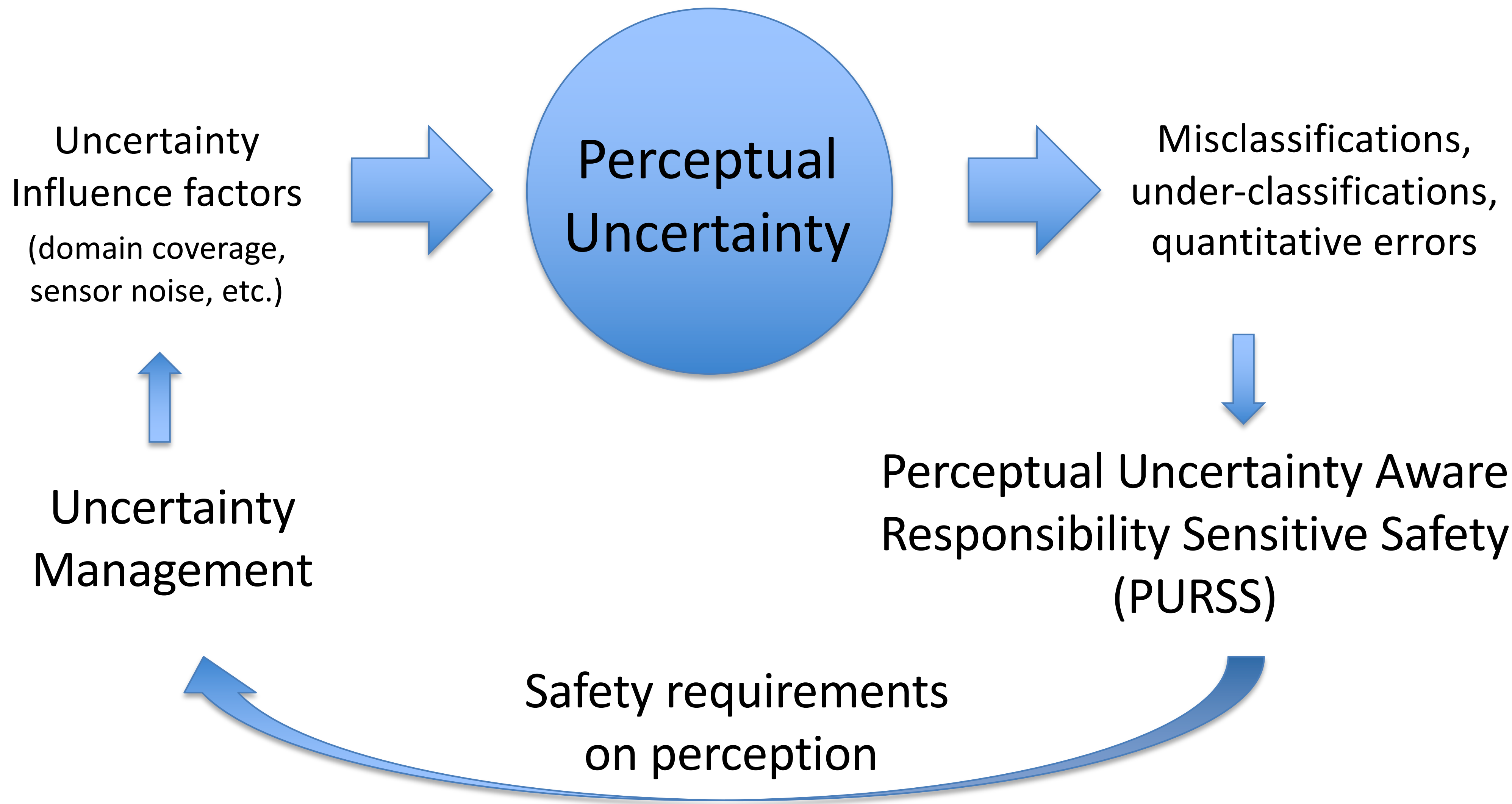
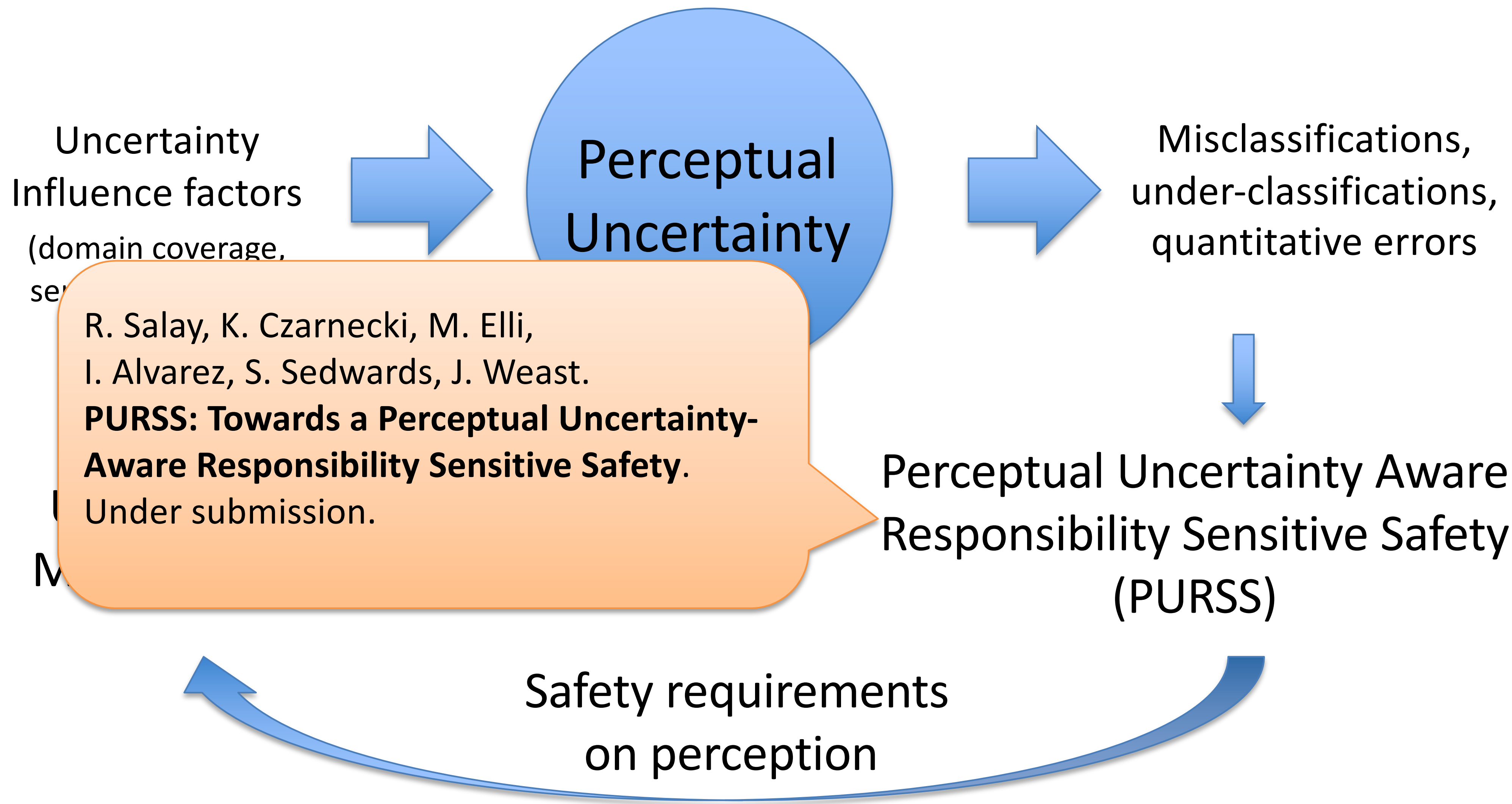
# Uncertainty-Centric Safety Assurance of ML-Based Perception for Automated Driving

Krzysztof Czarnecki
Waterloo Intelligent Systems Engineering (WISE) Lab
University of Waterloo

# Uncertainty-Centric Assurance of ML-Based Perception



# Uncertainty-Centric Assurance of ML-Based Perception



# Responsible Sensitive Safety (RSS)

- Defines responsible behavior to address **behavioral uncertainty**
  - Safe actions when safe and proper response when not safe
- Guarantees no collision when everyone follows the rules

# Responsible Sensitive Safety (RSS)

**RULE 1.** Do not hit the car in front
(longitudinal distance)

**RULE 2.** Do not cut in recklessly
(lateral distance)

**RULE 3.** Right of way is given, not taken

**RULE 4.** Be cautious in areas with limited visibility

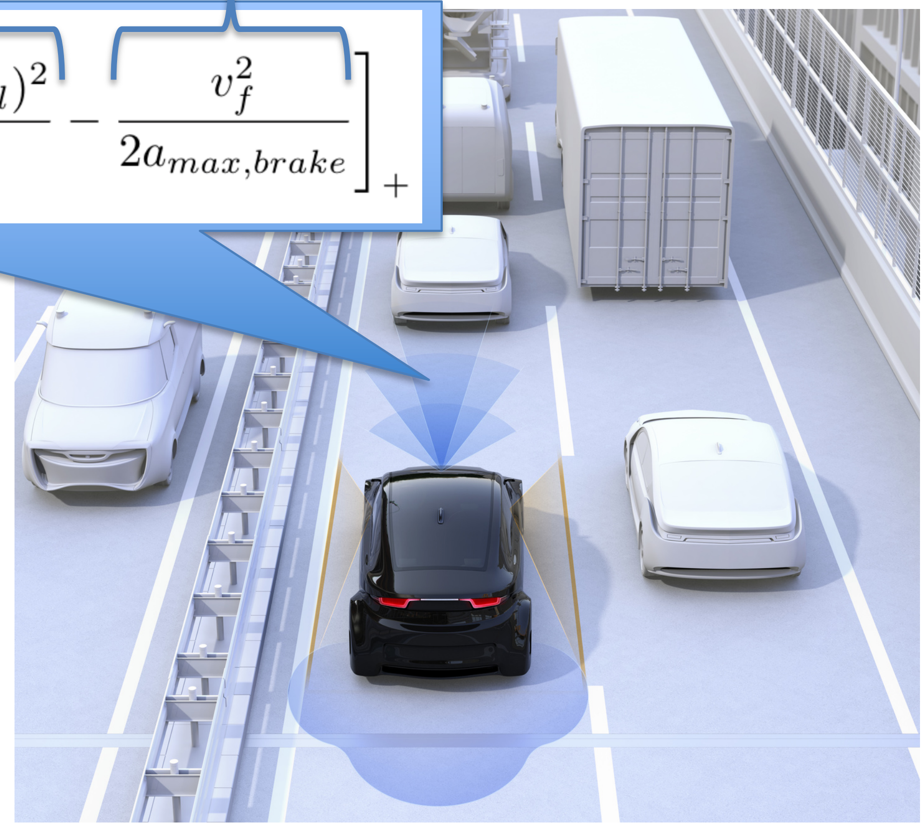
**RULE 5.** If you can avoid a crash without causing another one,
you must

https://arxiv.org/abs/1708.06374

# RULE 1. Safe Following Distance in RSS

Distance traveled due to reaction time — Braking distance — Distance traveled by front vehicle

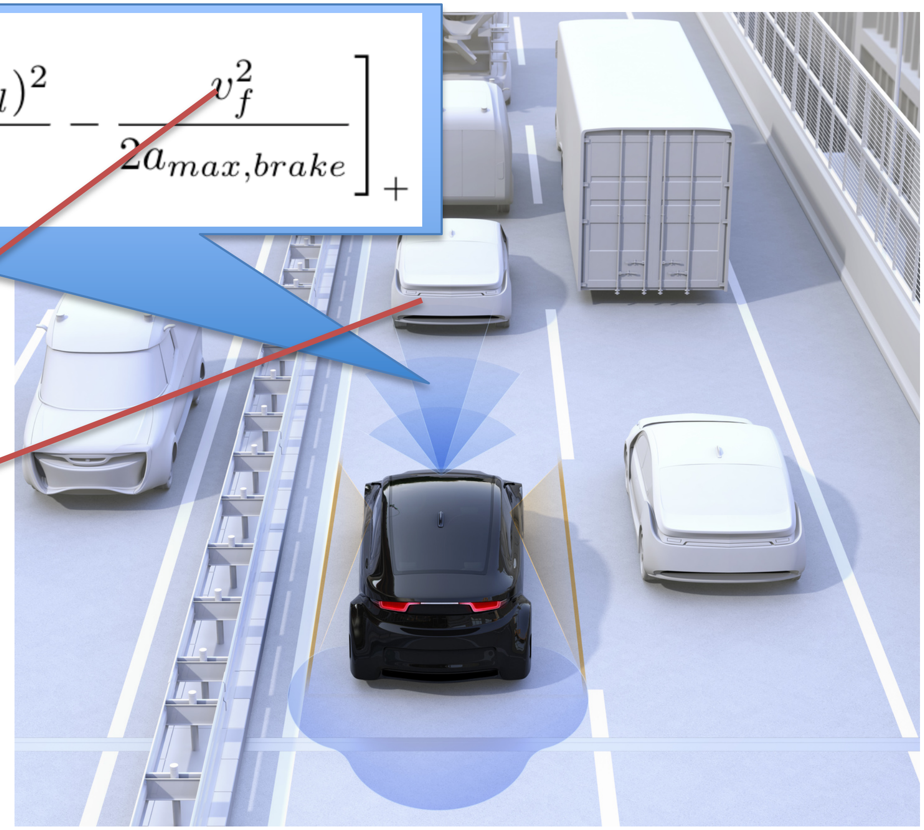
$$d_{min} = \left[ v_r\rho + \frac{1}{2}a_{max,accel}\rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$



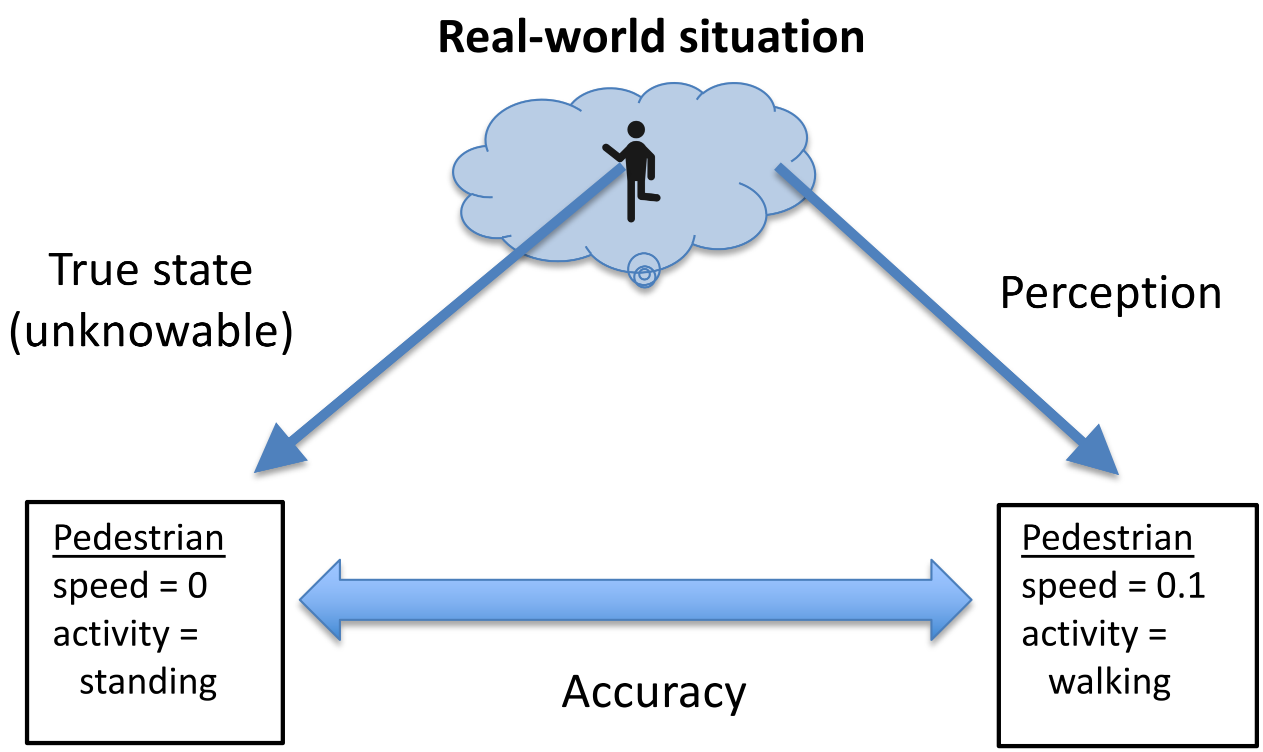
# RULE 1. Safe Following Distance in RSS

$$d_{min} = \left[ v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2 a_{min,brake}} - \frac{v_f^2}{2 a_{max,brake}} \right]_+$$

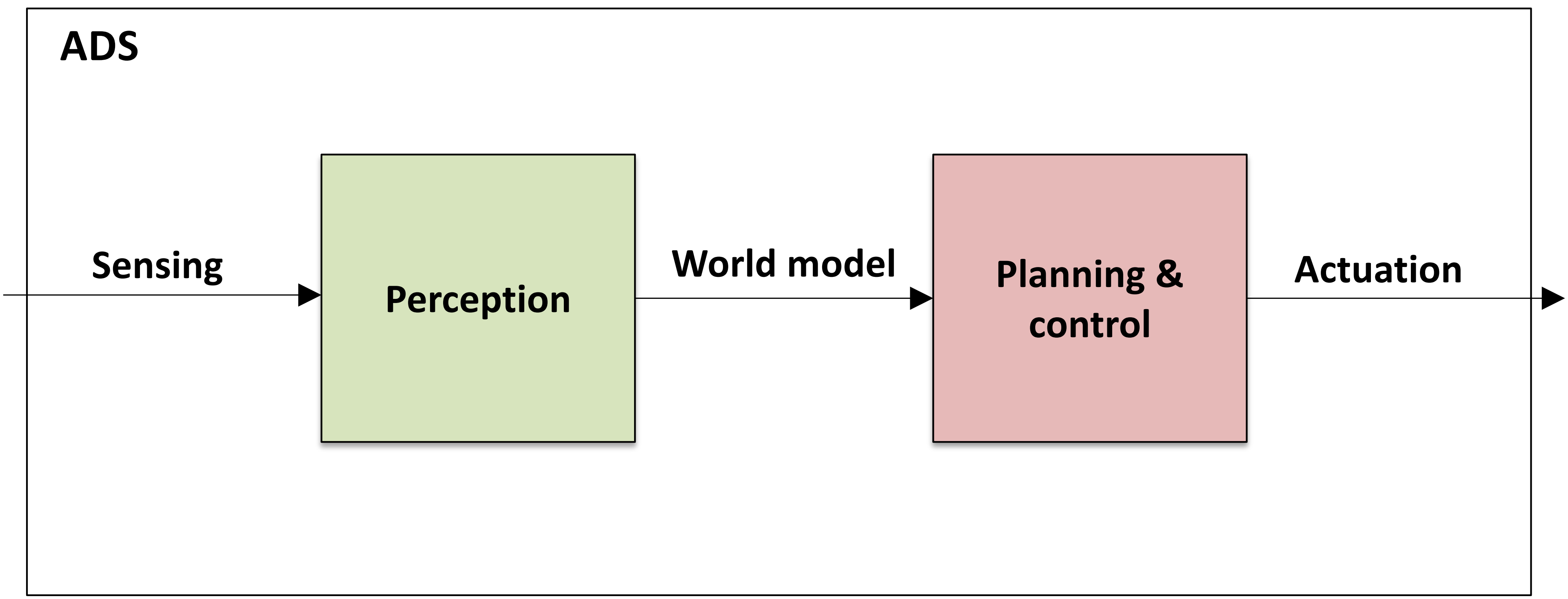
**Problem: Assumes perfect perception**



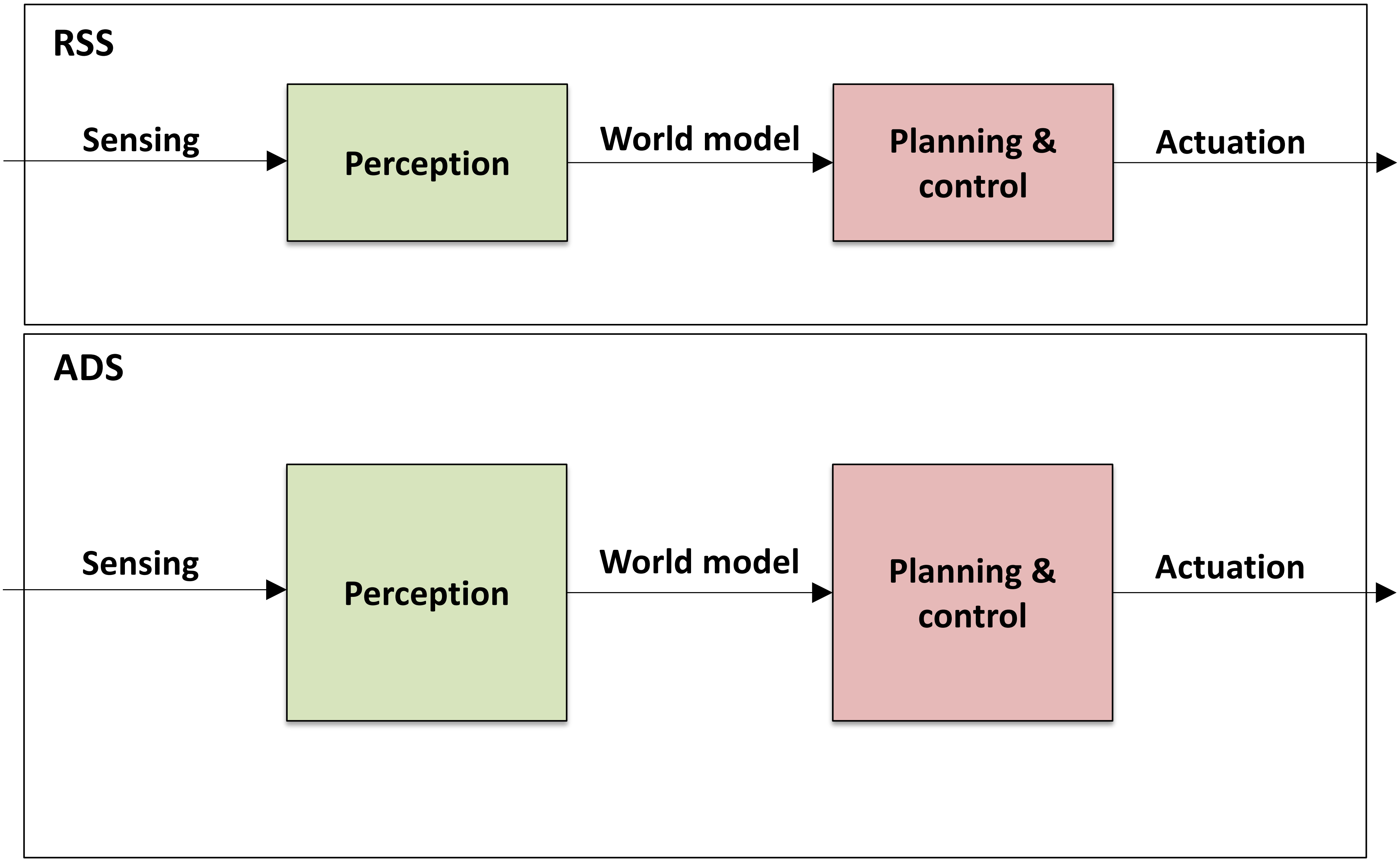
# Perception Triangle



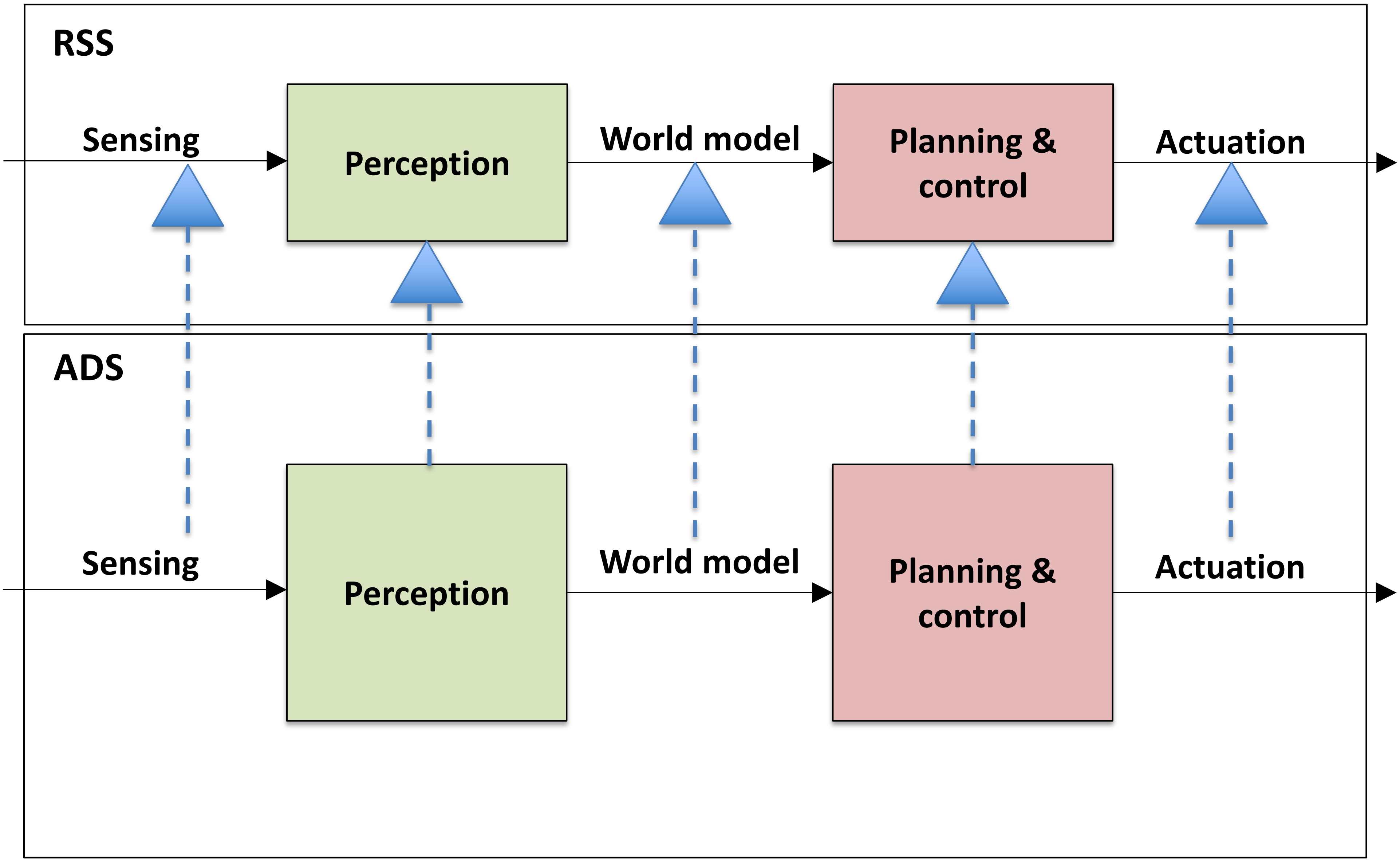
# Safety Argument Decomposition



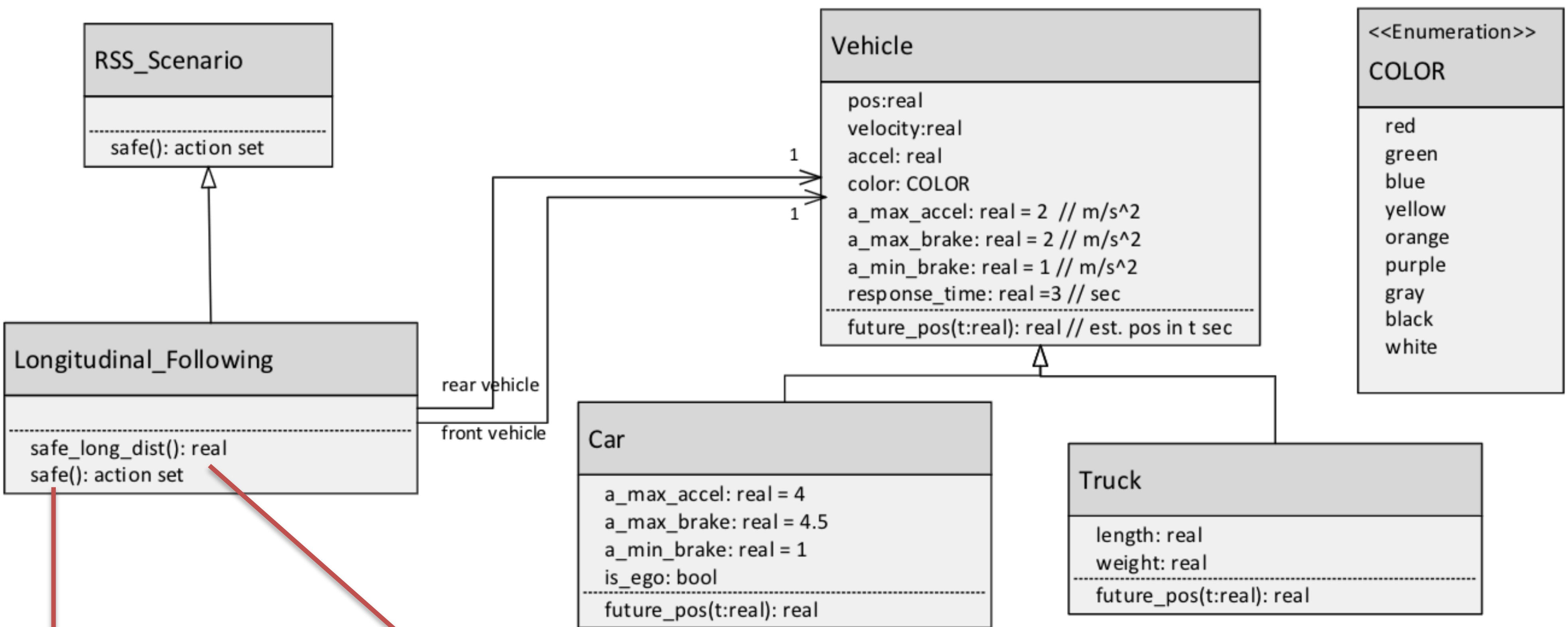
# RSS as a Constraint on ADS



# RSS as a Constraint on ADS



# Sample RSS-Compliant World Model Schema




**Safe following distance**

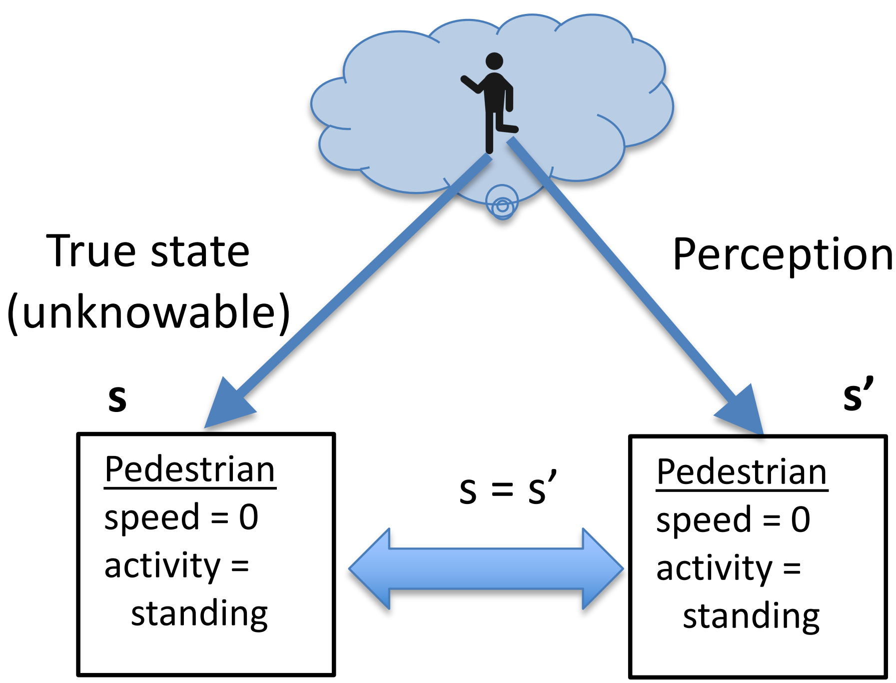
**Safe action set Safe(s)**

# Perception Cases (s → s')

## Correct Perception

s → s' where s = s'

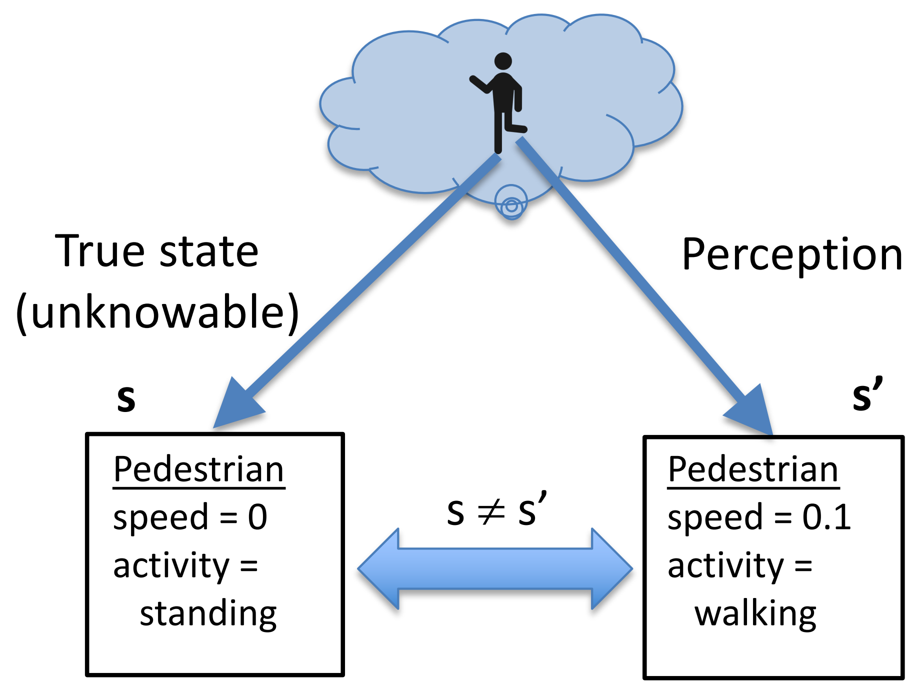
**Real-world situation**




## Misperception

s → s' where s ≠ s'

**Real-world situation**



# Safety of Perception

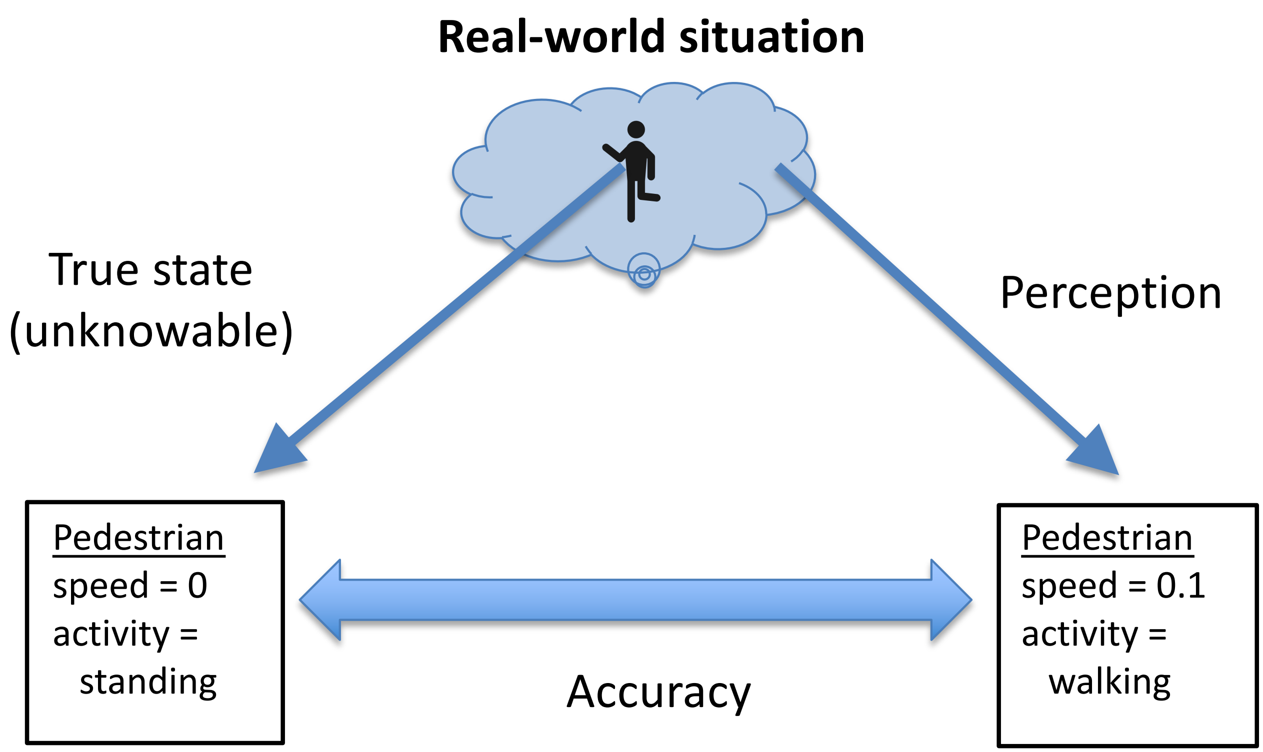
Misperception s → s’ potentially causes safety risk iff

$$\mathtt{Safe}(s') \not\subseteq \mathtt{Safe}(s).$$

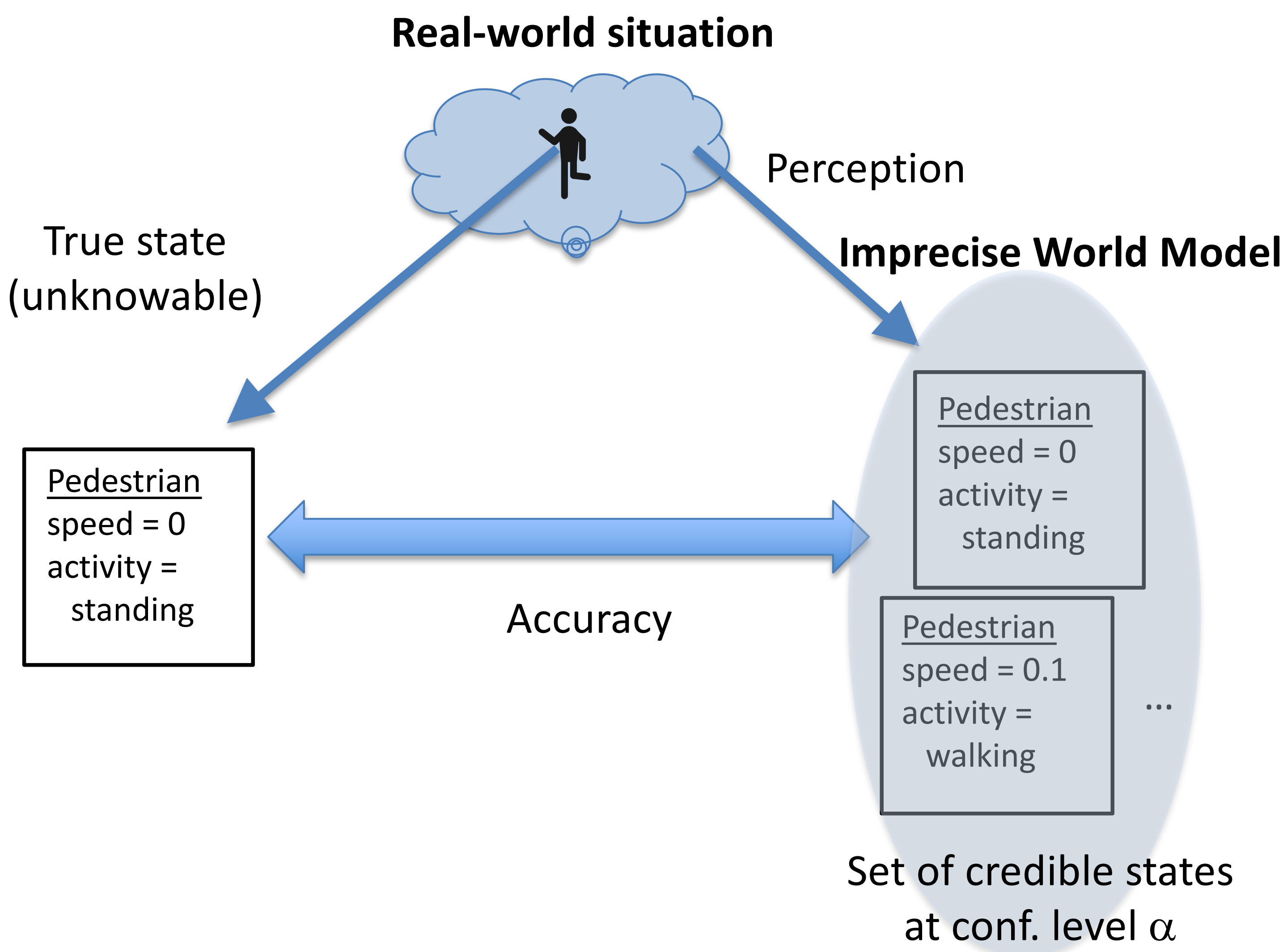
# Safety-Irrelevant Misperceptions

Misperception $s \rightarrow s'$ where $Safe(s) = Safe(s')$

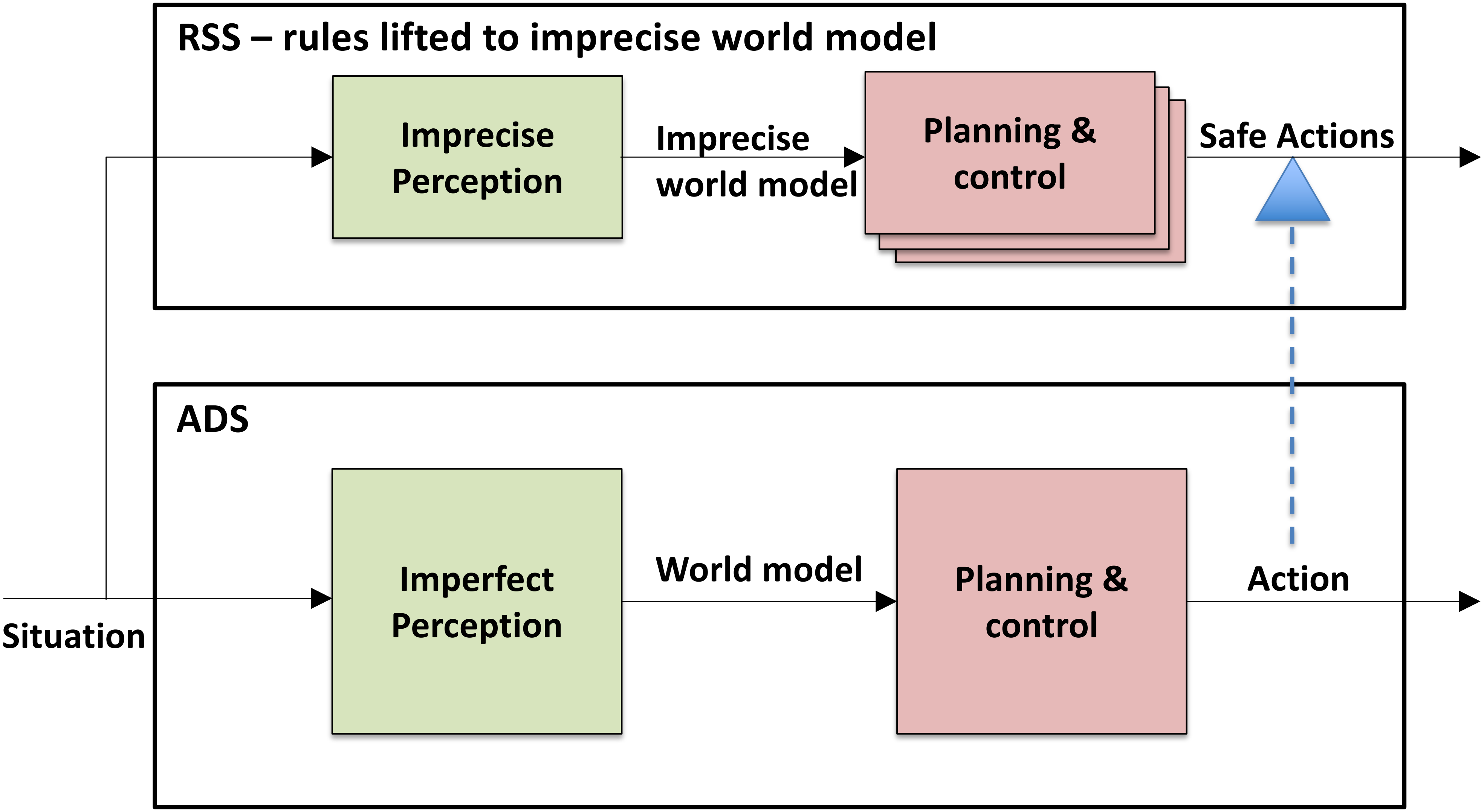
# Precise World Model



# Perceptual Uncertainty Handling via Imprecise World Models



# Perceptual Uncertainty Aware RSS (PURSS)



# Lifting World Model Schema to Imprecise World Model Schema

Elementwise lifting:

- Class entity to superclass
- Continuous value to interval
- Discrete value to enumerated set
- Derived attributes via set operations and interval arithmetic

# Using Imprecise World Models to Mitigate Misperception

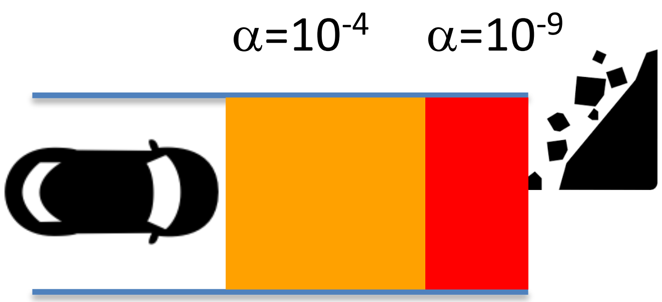
Given an under-perception case, where S is an imprecise model of confidence α perceived when the correct model:

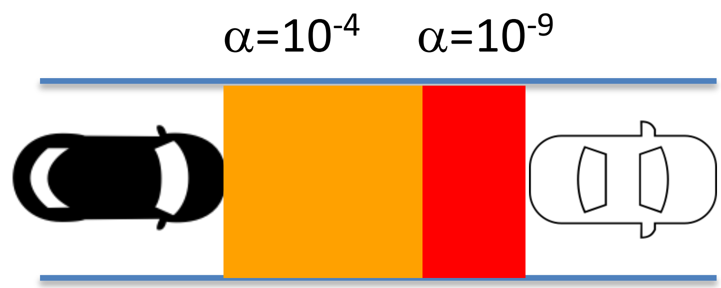
$$s \rightarrow_{\alpha} S$$

A safe action in an imprecise model must be safe for every precise model covered by the imprecise model.

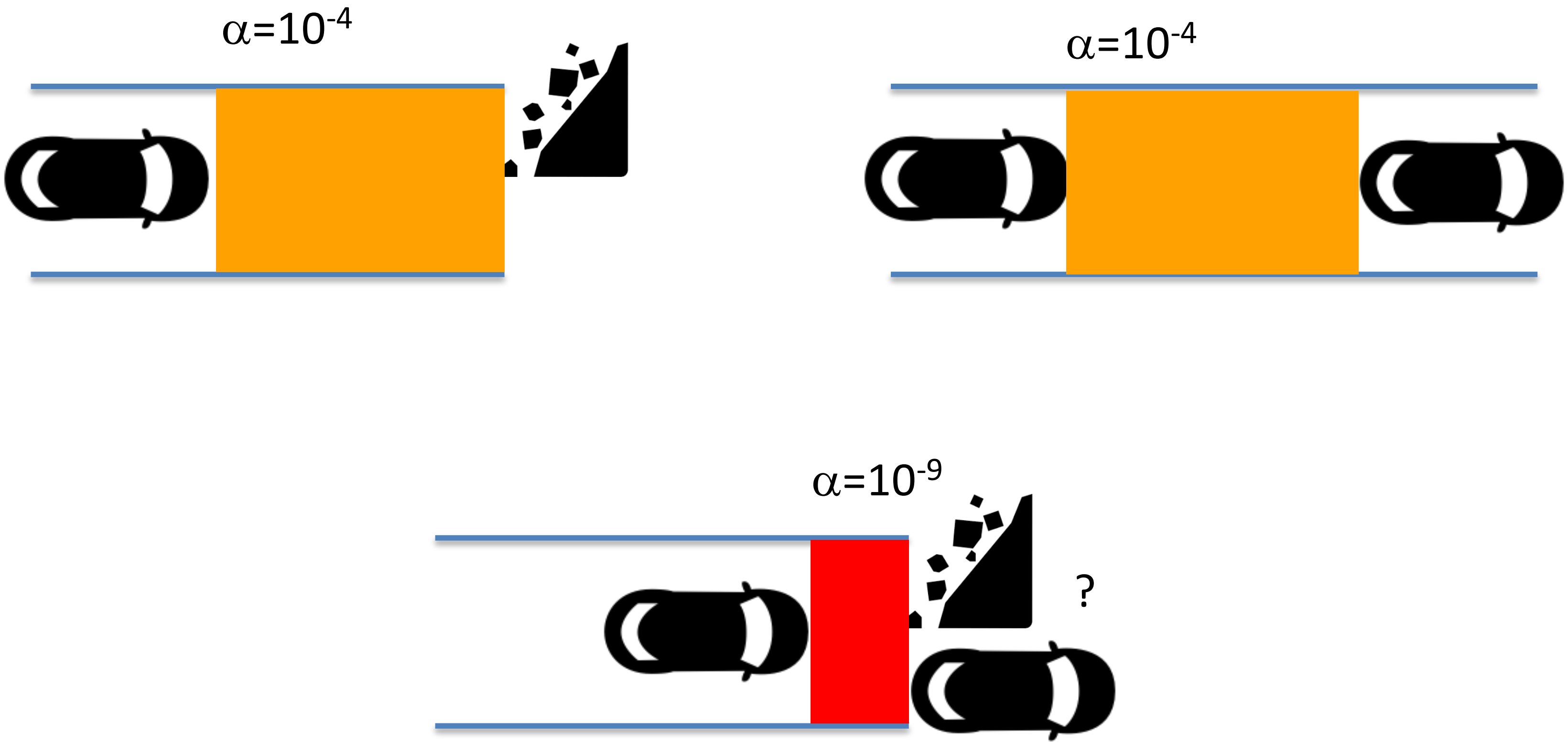
$$\texttt{Safe}(S) = \bigcap_{s_i \in S} \texttt{Safe}(s_i)$$

# Different Risk Levels

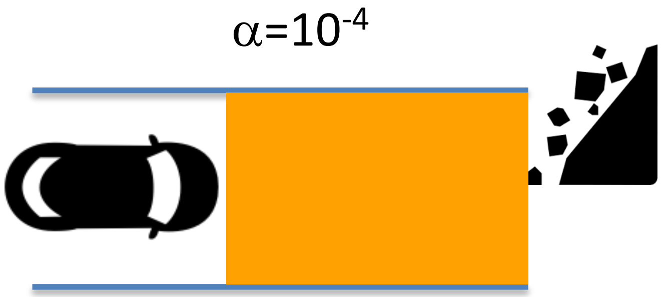


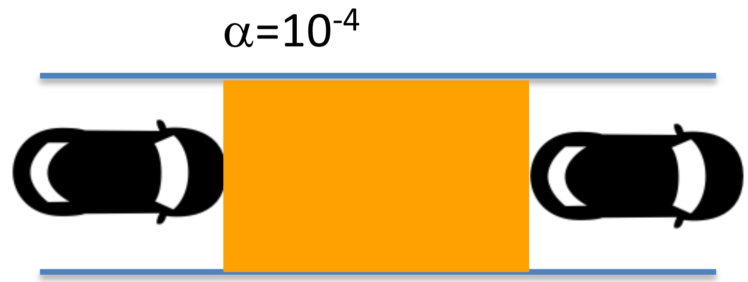



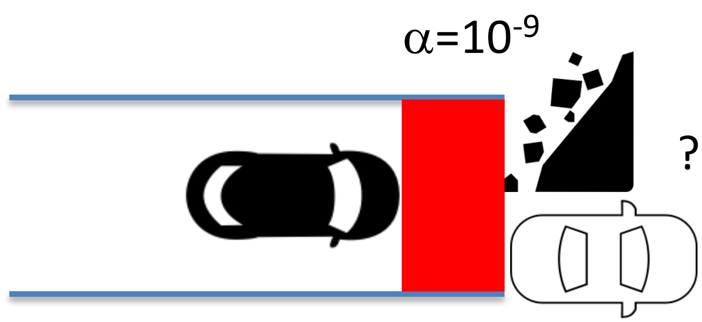
# Imprecise Classification when High Integrity Required



# Conservative Action for High Integrity

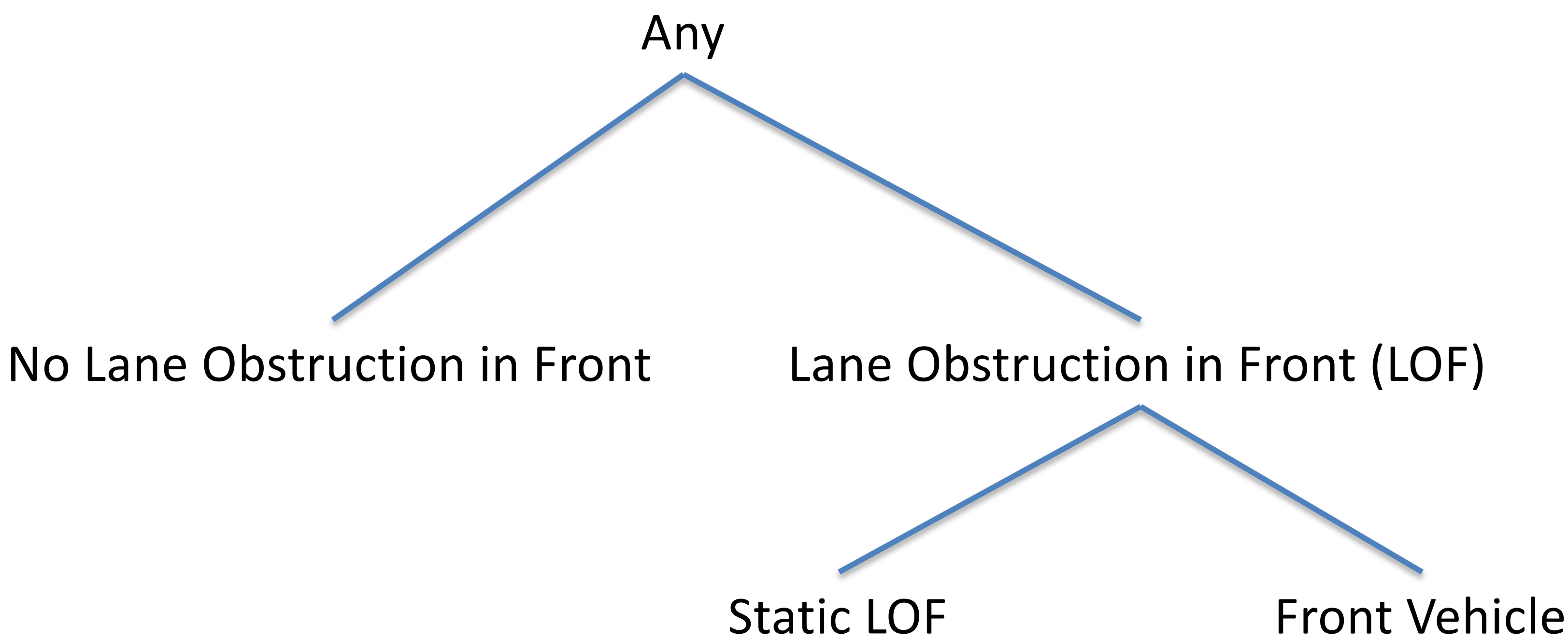


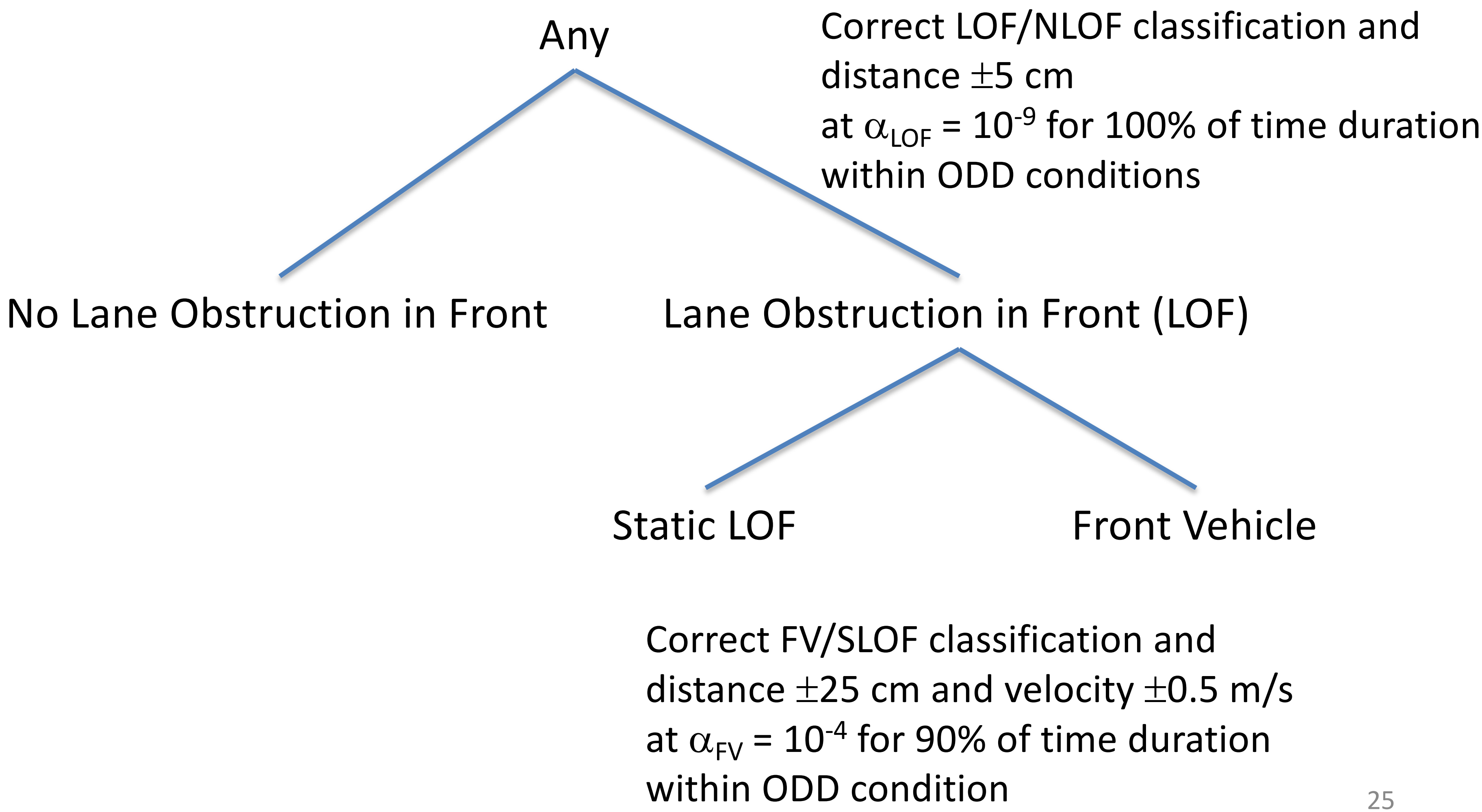

$$\mathtt{Safe}(S) = \bigcap_{s_i \in S} \mathtt{Safe}(s_i)$$

# Example of Mitigation

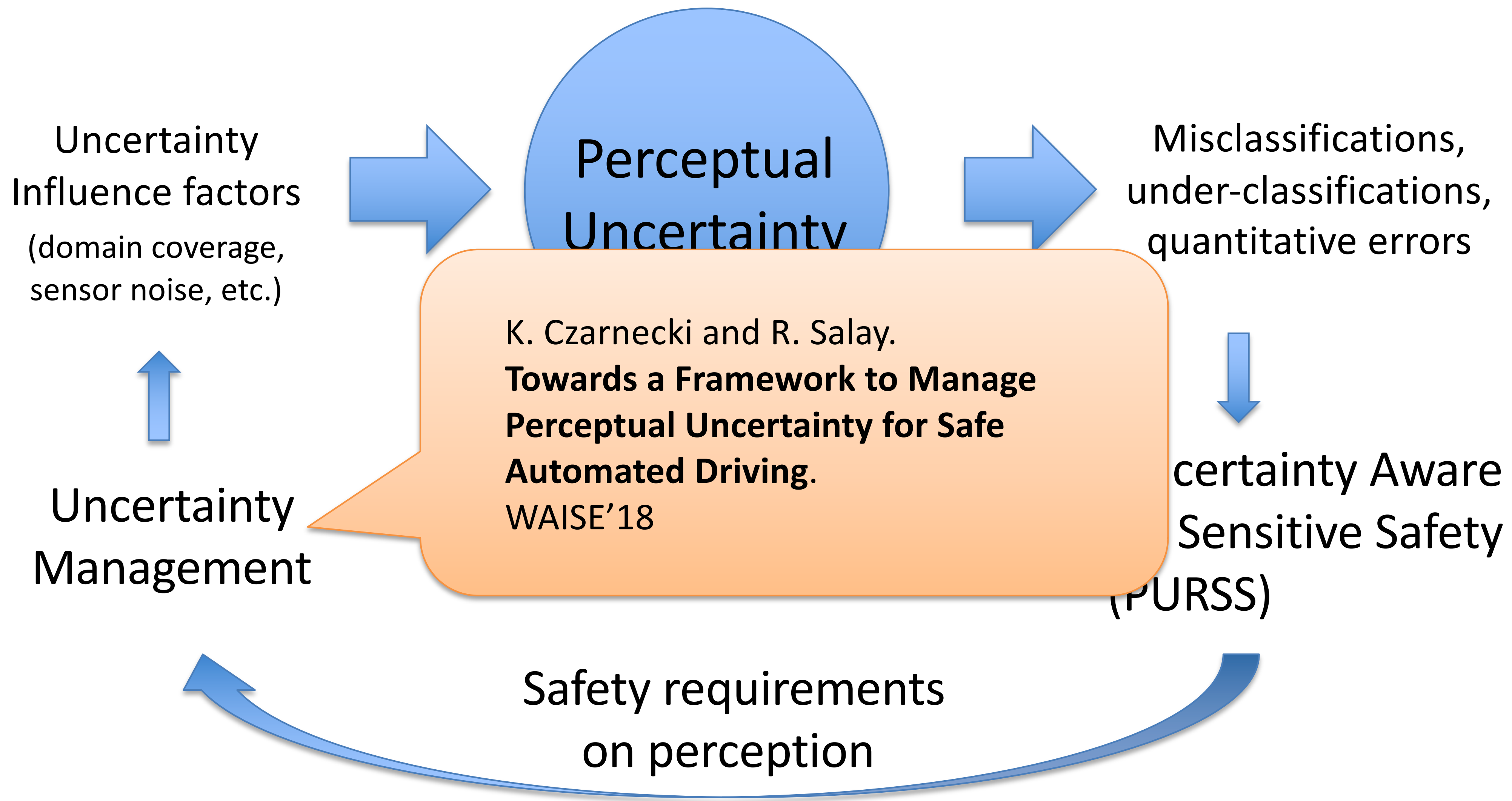



**Actions: continue or stop or follow**

# Safety Requirements on Perception Performance from PURSS

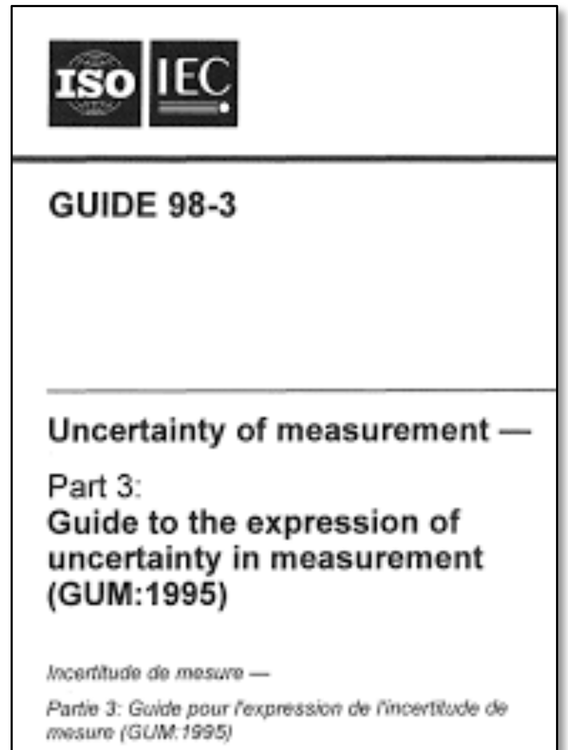


# Uncertainty-Centric Assurance of ML-Based Perception



# Guide to the Expression of Uncertainty in Measurement (GUM)

- True accuracy unknowable
  - Accuracy in ML wrt. test set only
- Must estimate uncertainty


ISO IEC

GUIDE 98-3

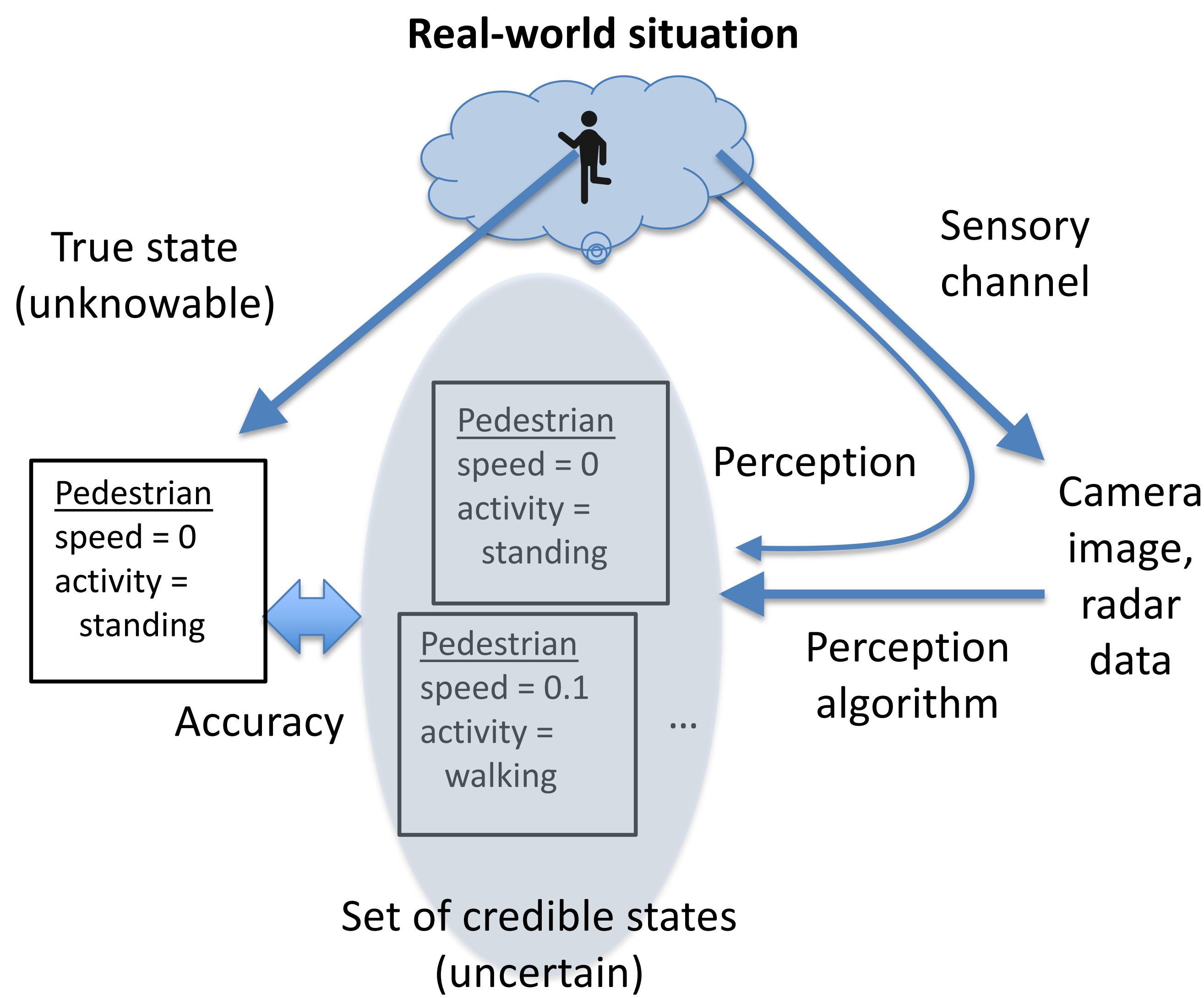
**Uncertainty of measurement —**

Part 3:
**Guide to the expression of uncertainty in measurement (GUM:1995)**

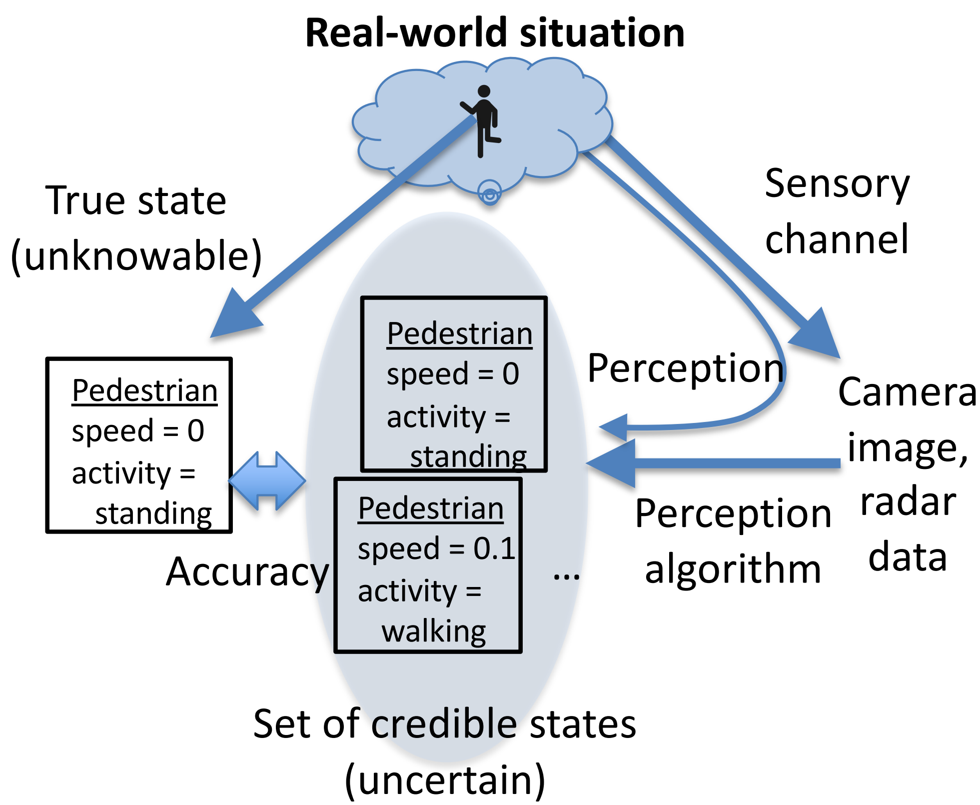
*Incertitude de mesure —*

*Partie 3: Guide pour l'expression de l'incertitude de mesure (GUM:1995)*

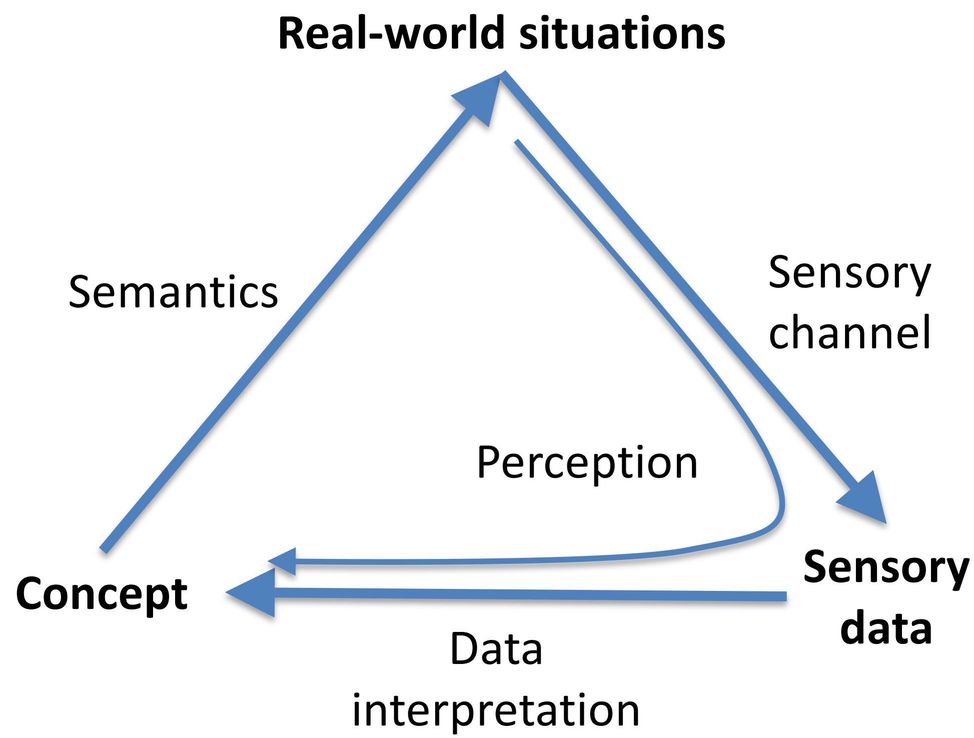
# Perception Triangle (Instance-Level)



# Perceptual Triangle

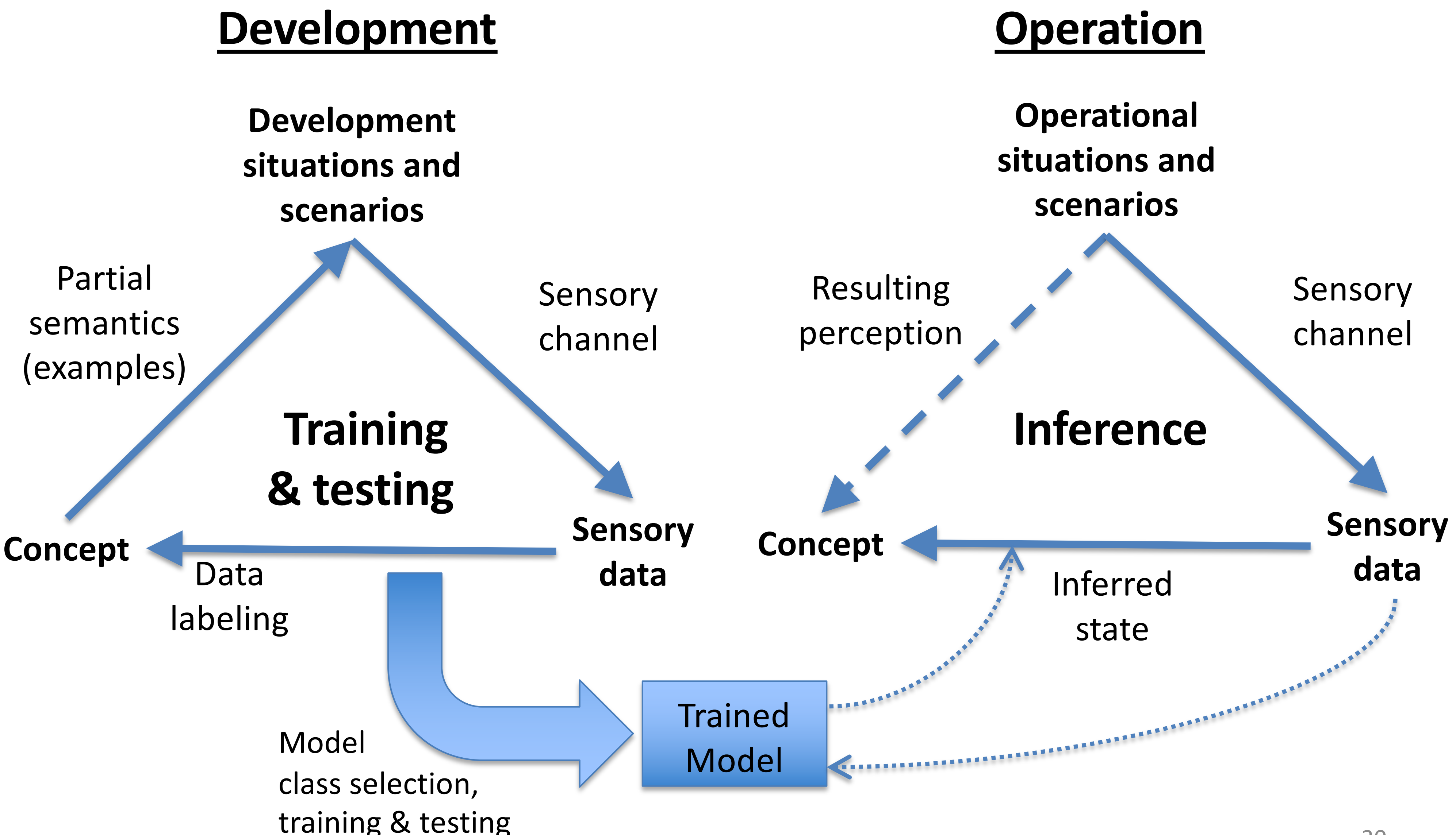



**Instance-level**

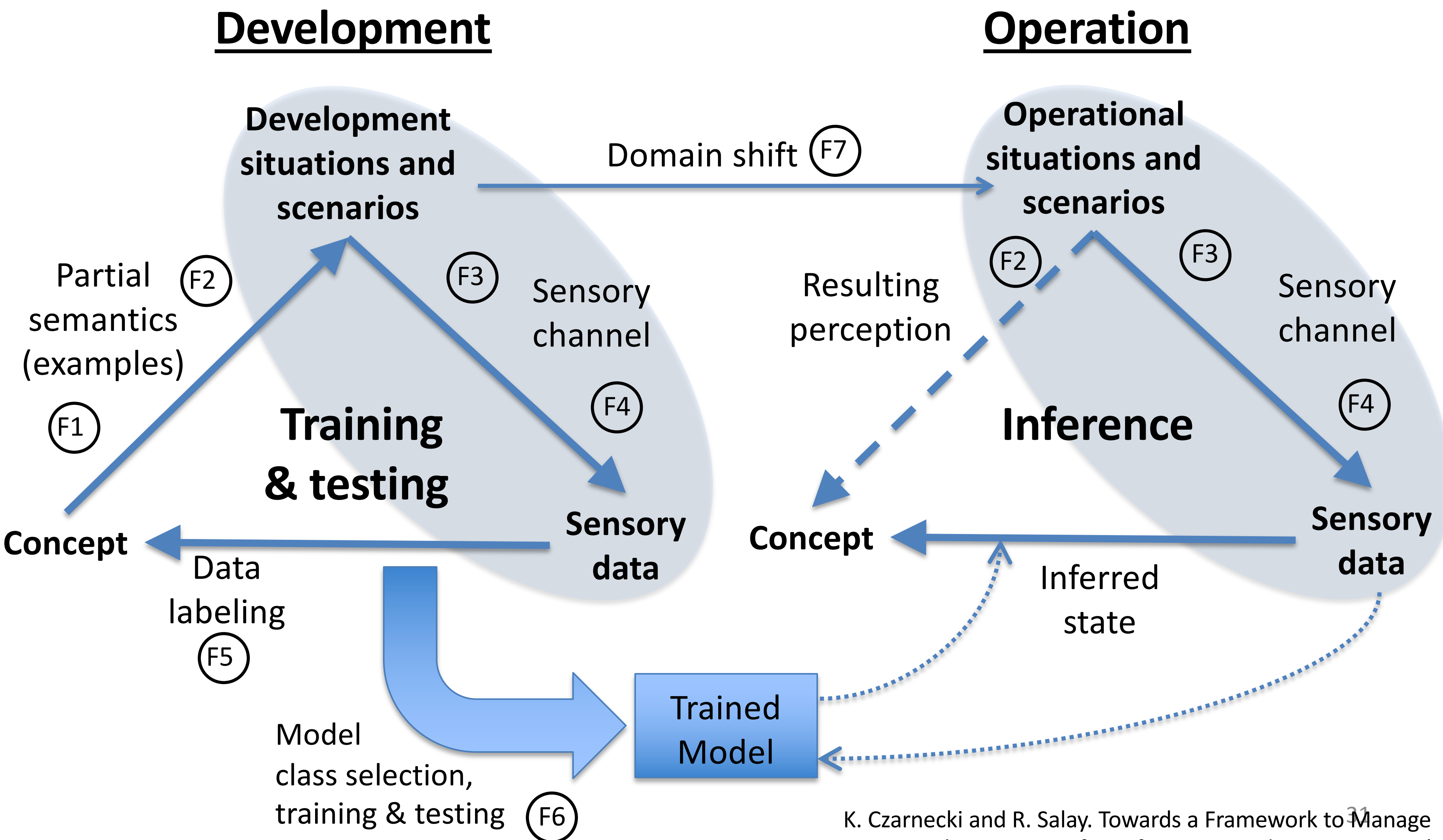



**Domain-level (generic)**

# Perceptual Triangle When Using Supervised ML

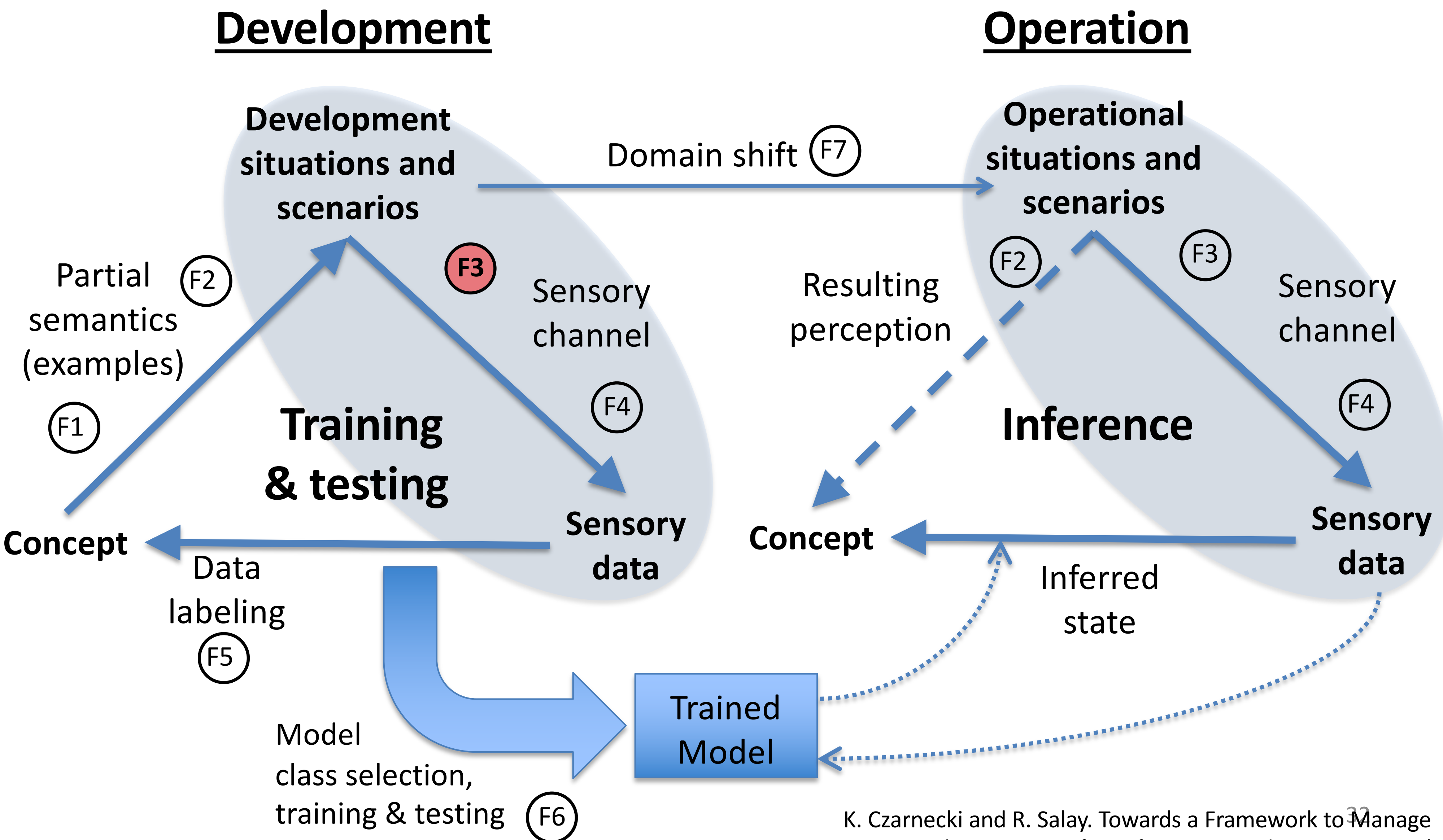


# Factors Influencing Uncertainty (F1-7)



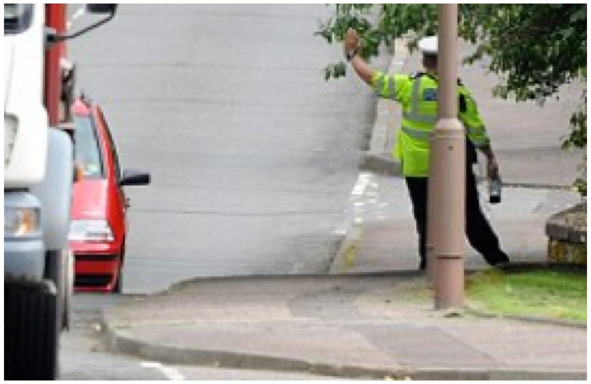

K. Czarnecki and R. Salay. Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving. WAISE'18

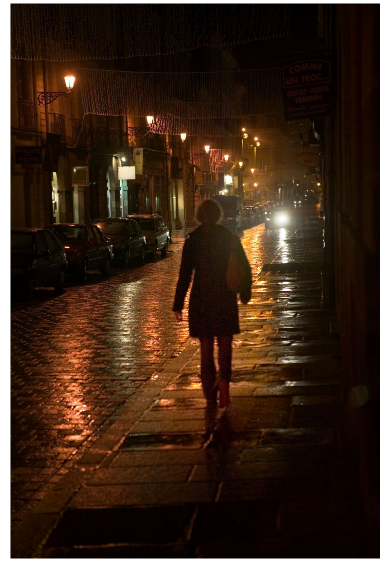
# Factors Influencing Uncertainty (F1-7)




K. Czarnecki and R. Salay. Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving. WAISE'18

# F3: Scene Uncertainty





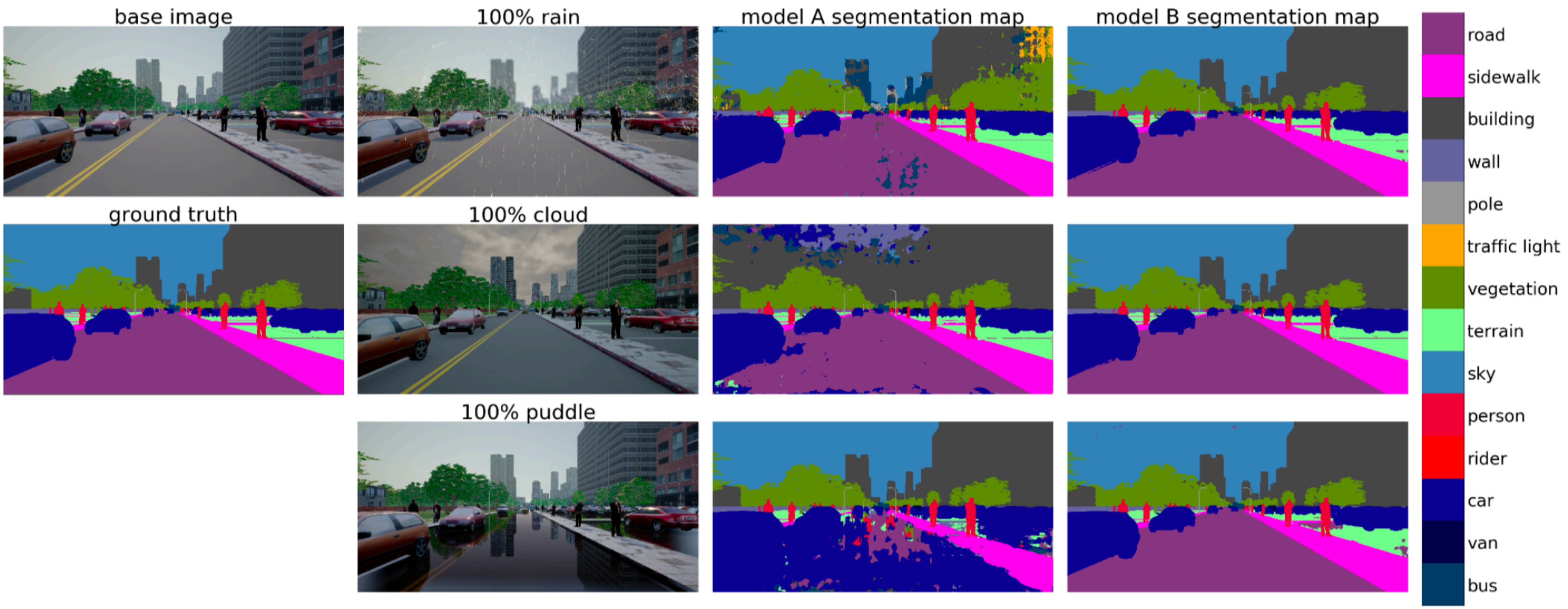
# F3: Scene Uncertainty

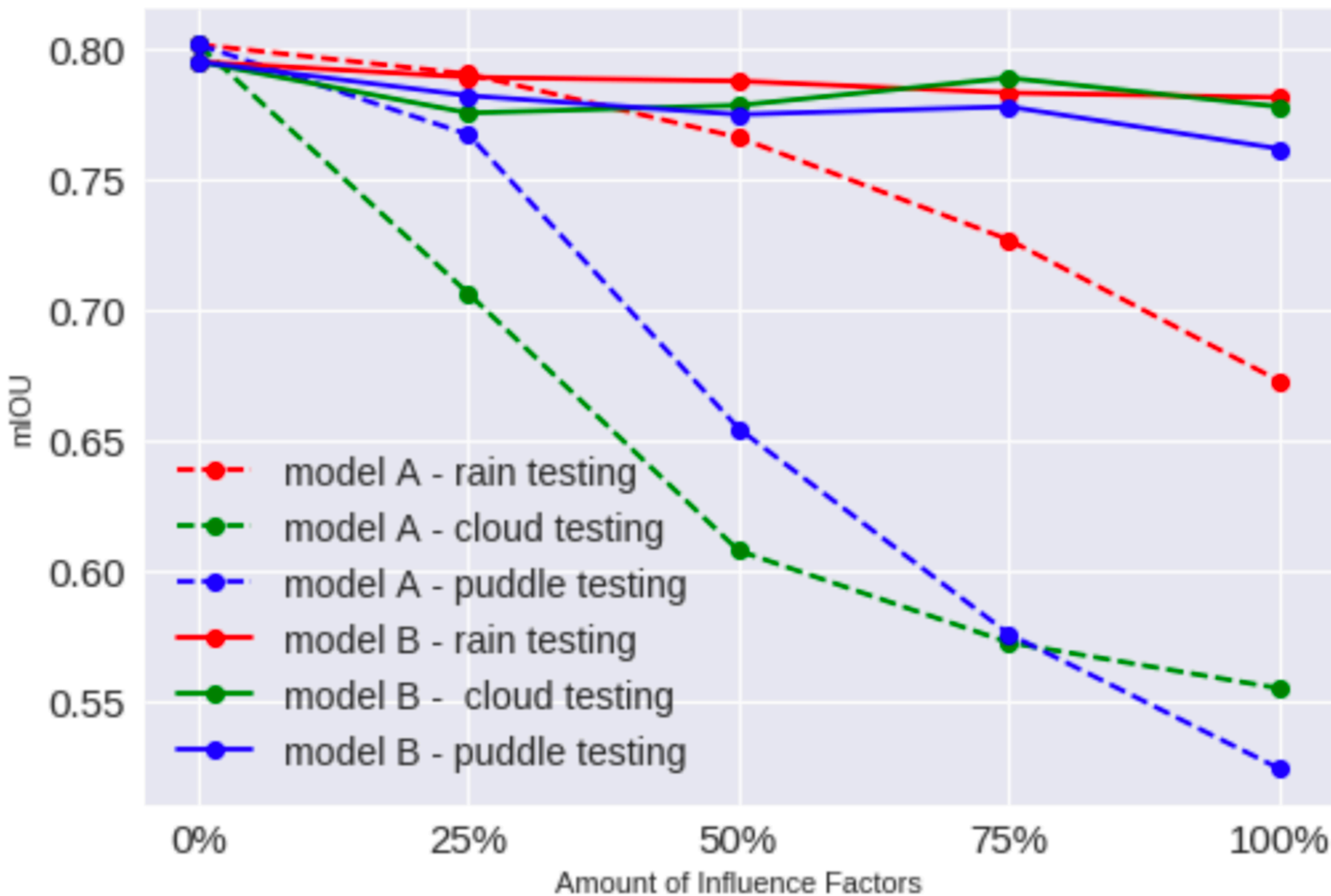
- Surrogate measures
  - range, scale, occlusion level, atmospheric visibility, illumination, clutter and crowding level
- Also part of development data set coverage
- To determine sufficient coverage, compare these measures with
  1. Test set accuracy
  2. Estimated uncertainty by the network

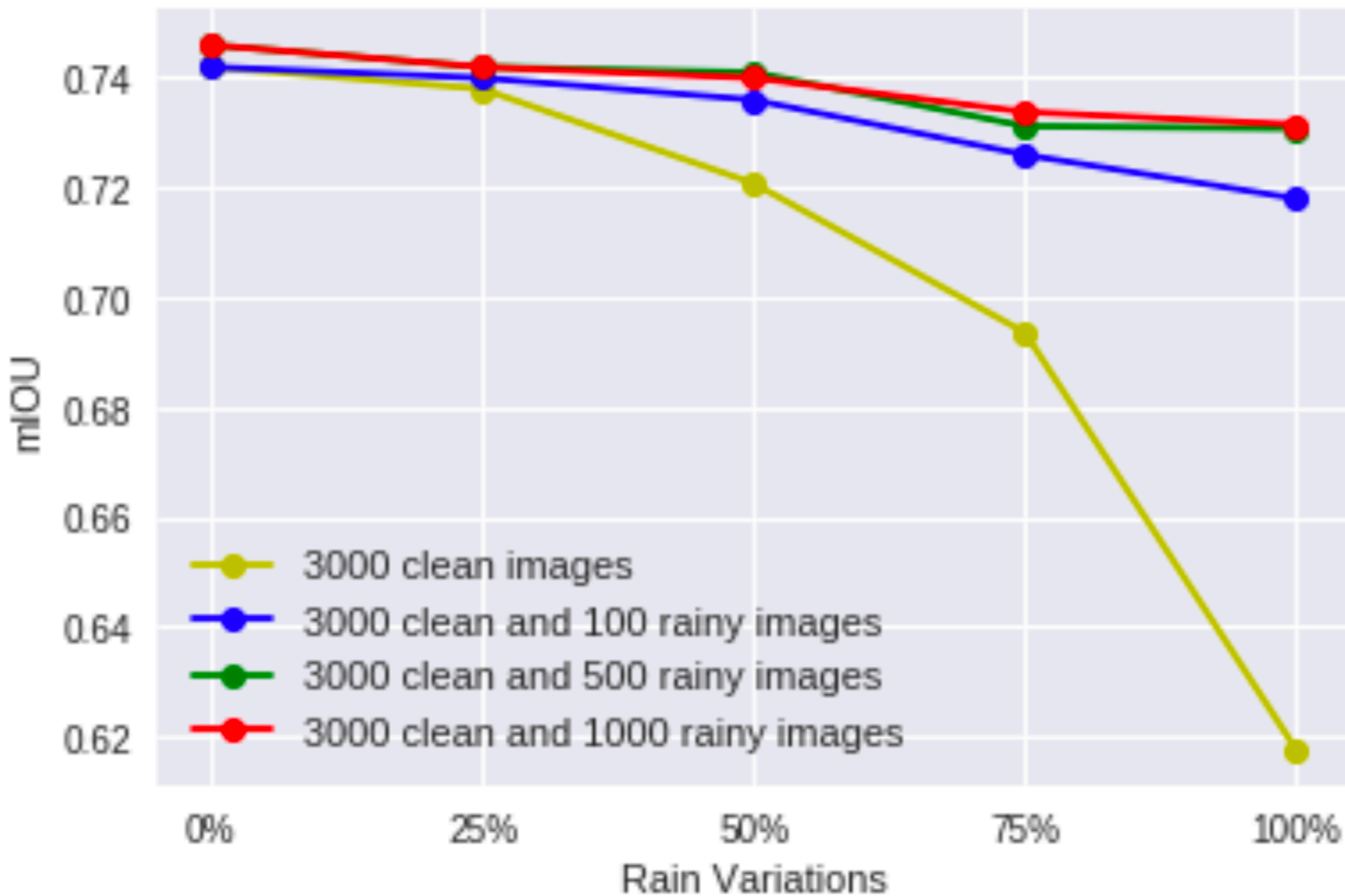
# Synthetic Dataset to Study Scene Influence Factors


Samin Khan, Buu Phan, Rick Salay, and Krzysztof Czarnecki. ProcSy: Procedural Synthetic Dataset Generation Towards Influence Factor Studies Of Semantic Segmentation Networks. Workshop on Vision for All Seasons: Bad Weather and Nighttime, associated with CVPR, Long Beach, 2019

# Scene Influence Factors -> Accuracy




0.80
0.75
0.70
0.65
0.60
0.55
mIOU
model A - rain testing
model A - cloud testing
model A - puddle testing
model B - rain testing
model B - cloud testing
model B - puddle testing
0%
25%
50%
75%
100%
Amount of Influence Factors



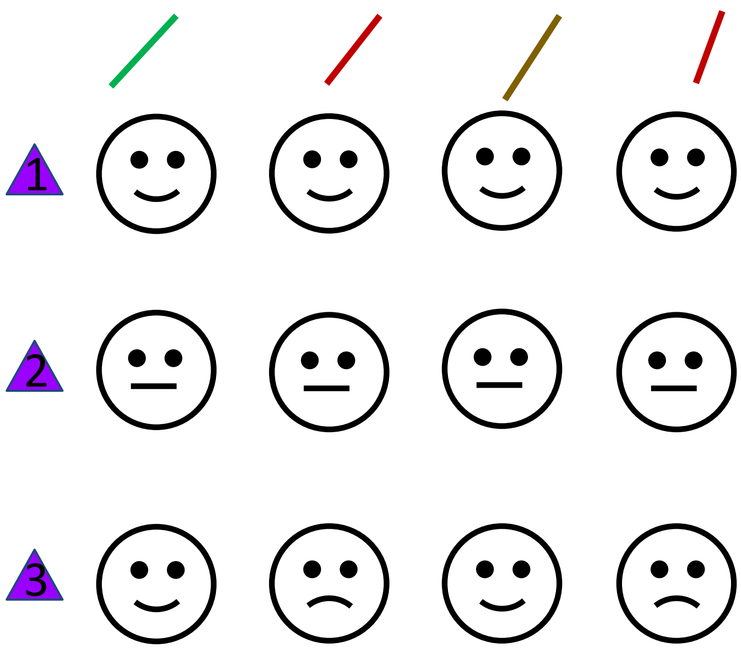
0.74
0.72
0.70
0.68
0.66
0.64
0.62
mIOU
0%
25%
50%
75%
100%
Rain Variations
3000 clean images
3000 clean and 100 rainy images
3000 clean and 500 rainy images
3000 clean and 1000 rainy images

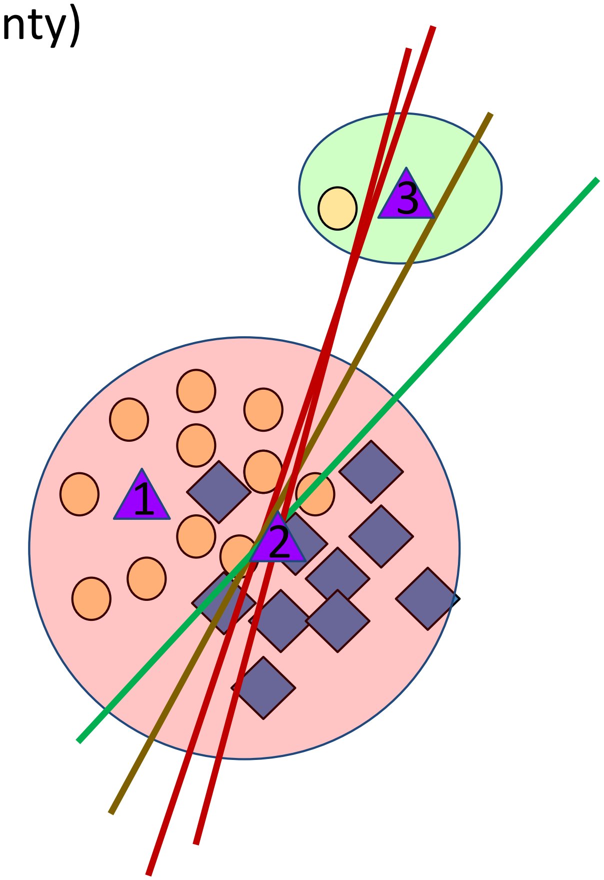
# Aleatoric and Epistemic Uncertainty

Predictive Entropy (PE) = H(E(p))

Aleatoric Entropy (AE) = E(H(p))

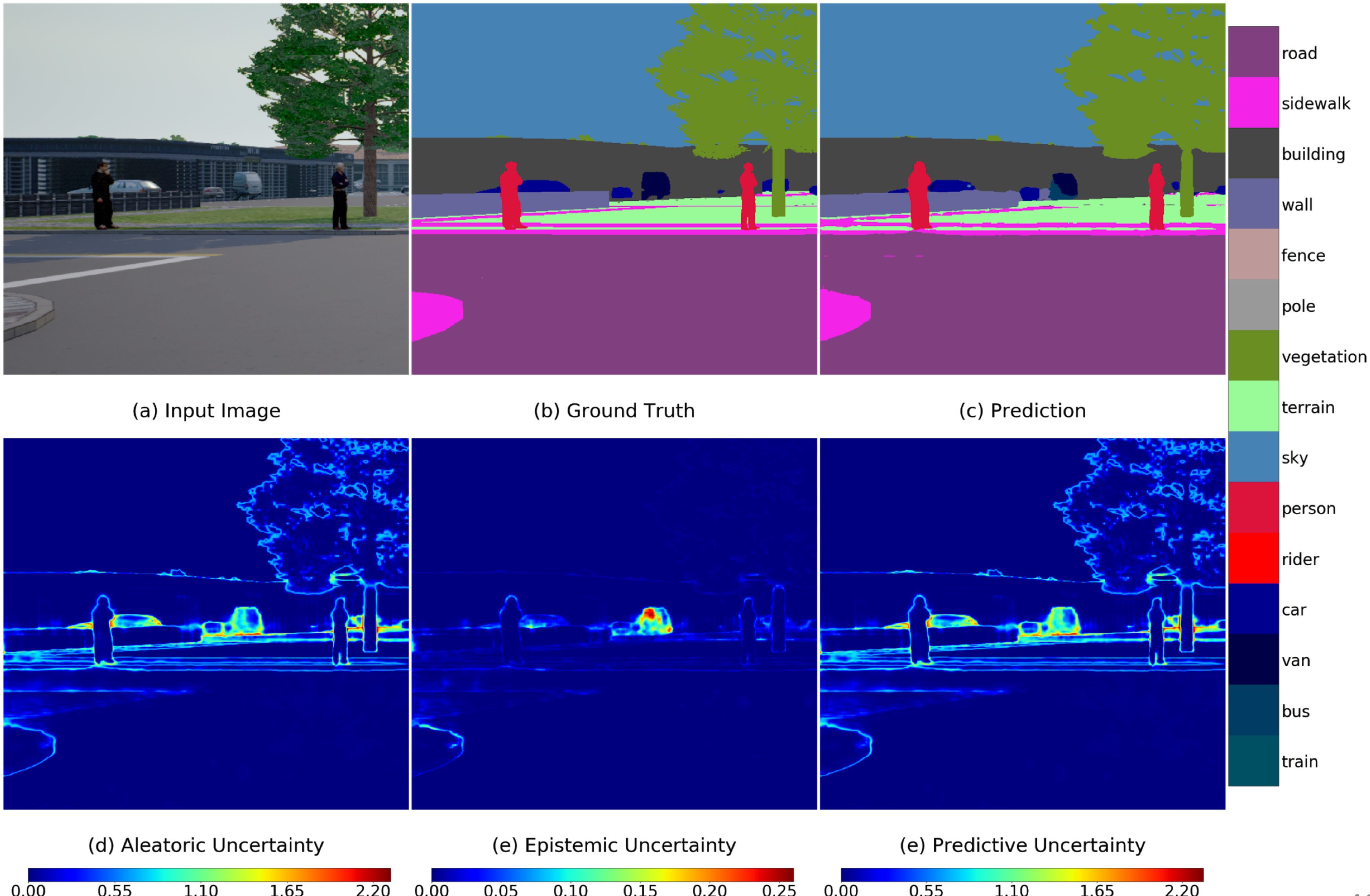
Mutual Information (MI) = PE - AE (Epistemic Uncertainty)



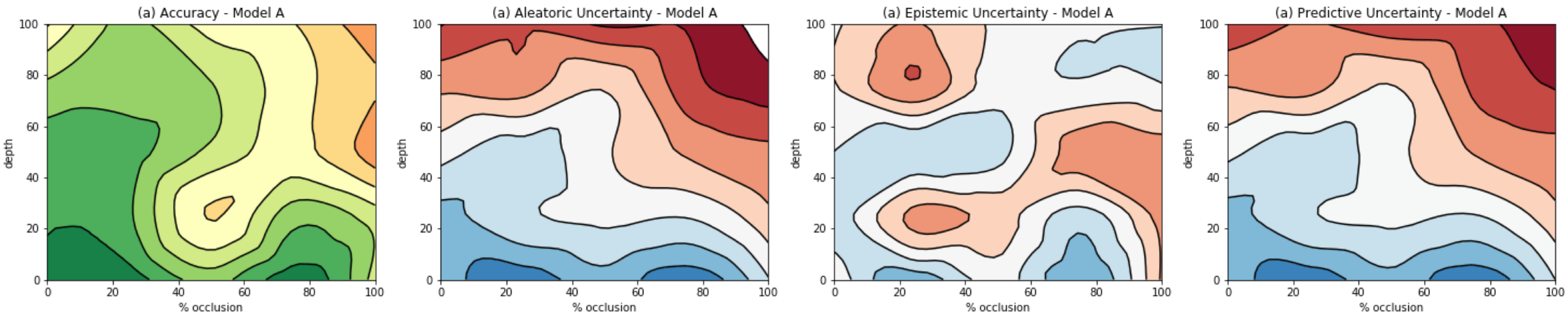




Smith L, Gal Y. Understanding measures of uncertainty for adversarial example detection. arXiv preprint arXiv:1803.08533

# Scene Influence Factors -> Uncertainty Estimates

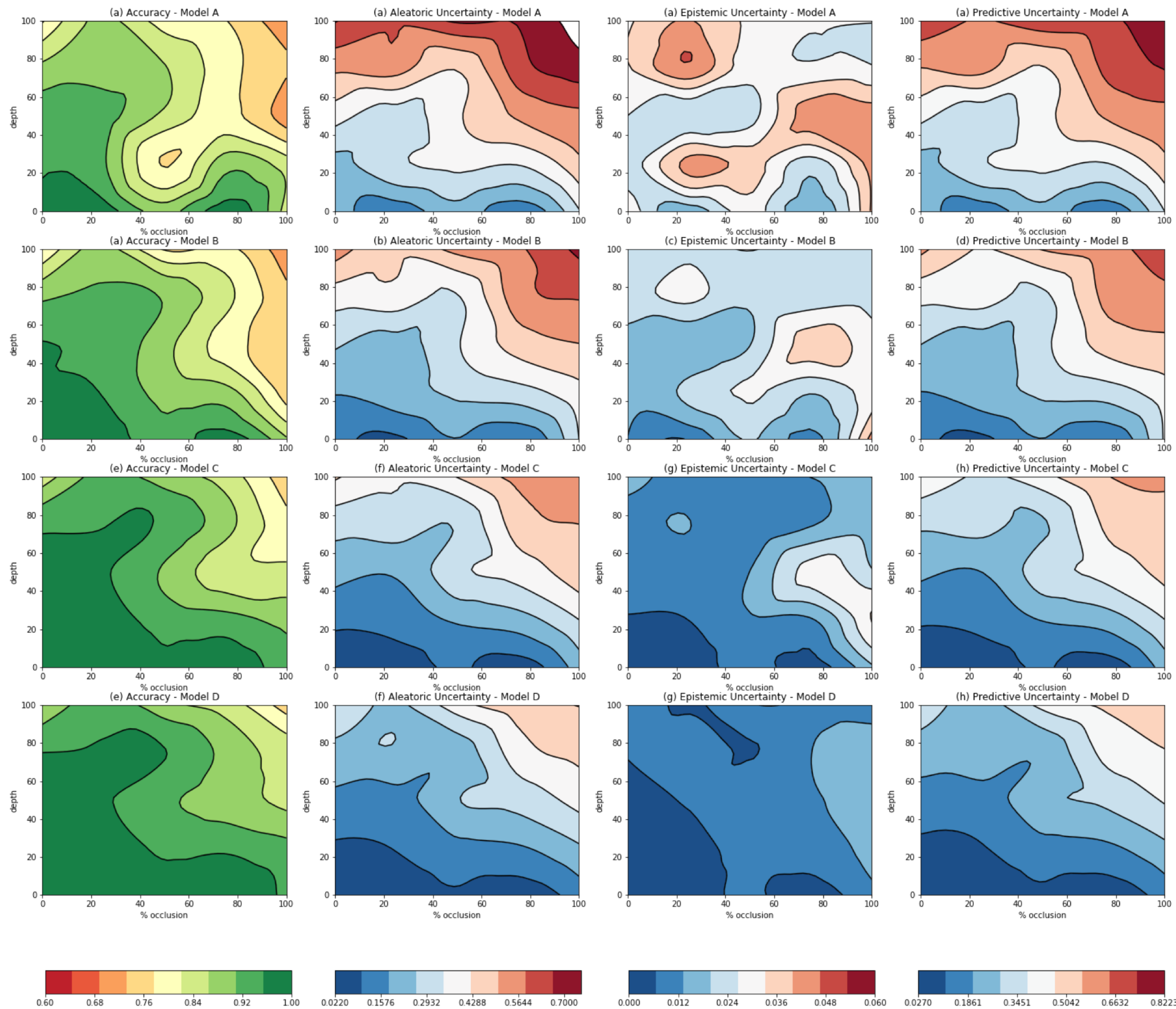


# Occlusion and Depth -> Uncertainty Estimates




Buu Phan, Samin Khan, and Rick Salay, and Krzysztof Czarnecki. Bayesian Uncertainty Quantification with Synthetic Data. In Proceedings of International Workshop on Artificial Intelligence Safety Engineering (WAISE), SAFECOMP, Turku, Finland, 2019

# Occlusion and Depth -> Uncertainty Estimates



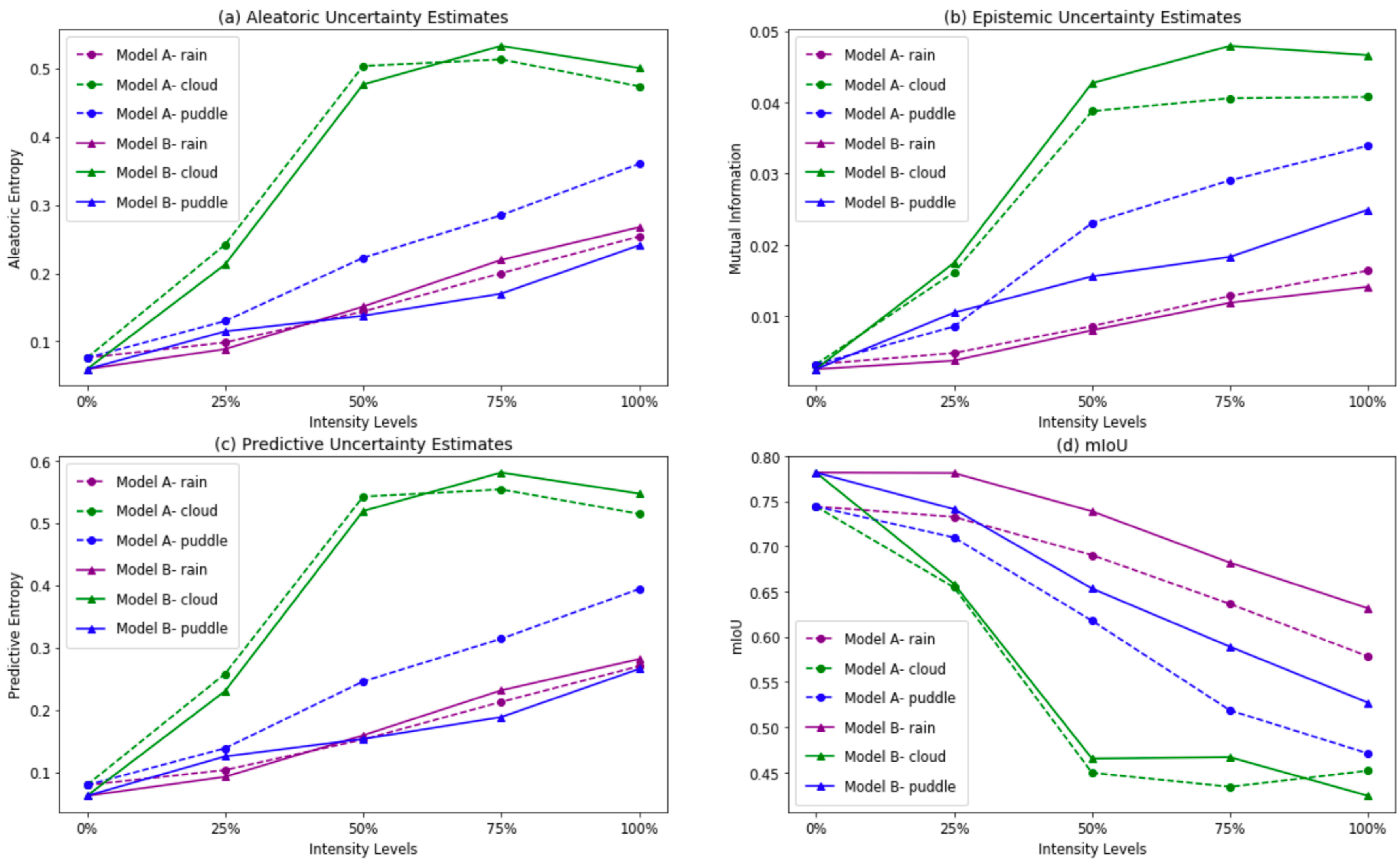

|D| = 500

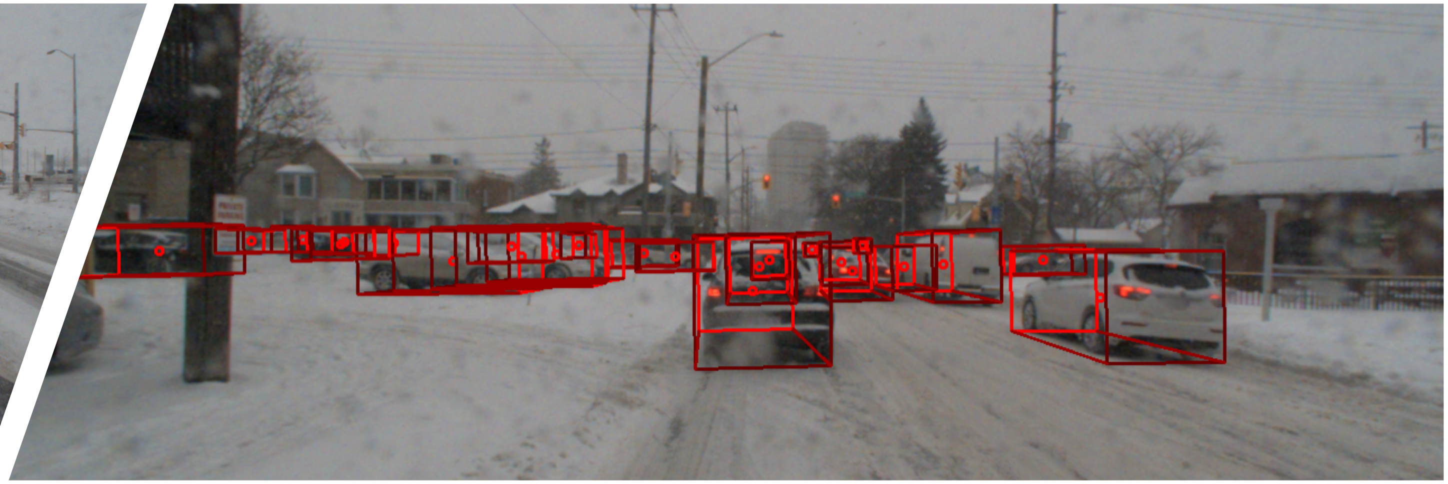
|D| = 3000

|D| = 8000

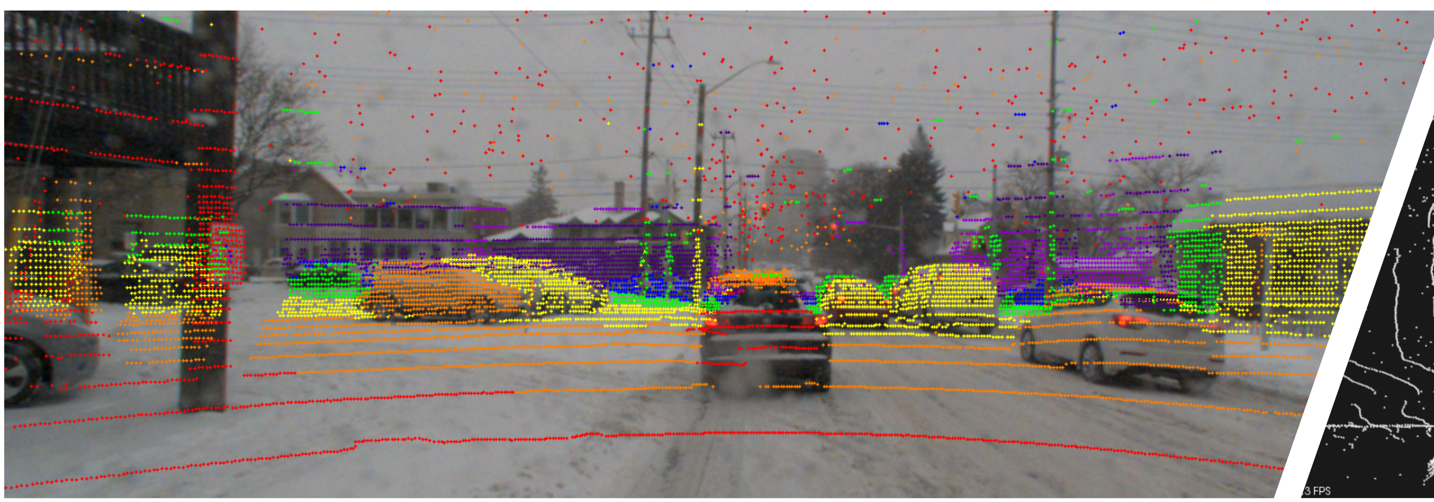
|D| = 13100

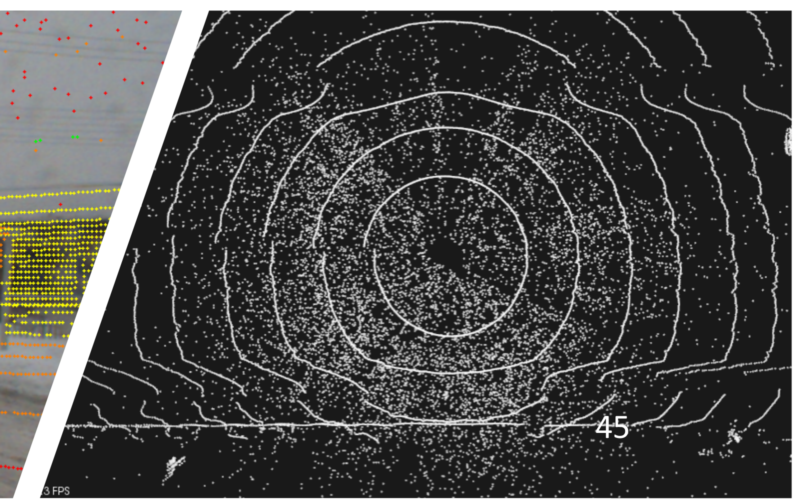
# Rain, Clouds, Puddles -> Uncertainty Estimates



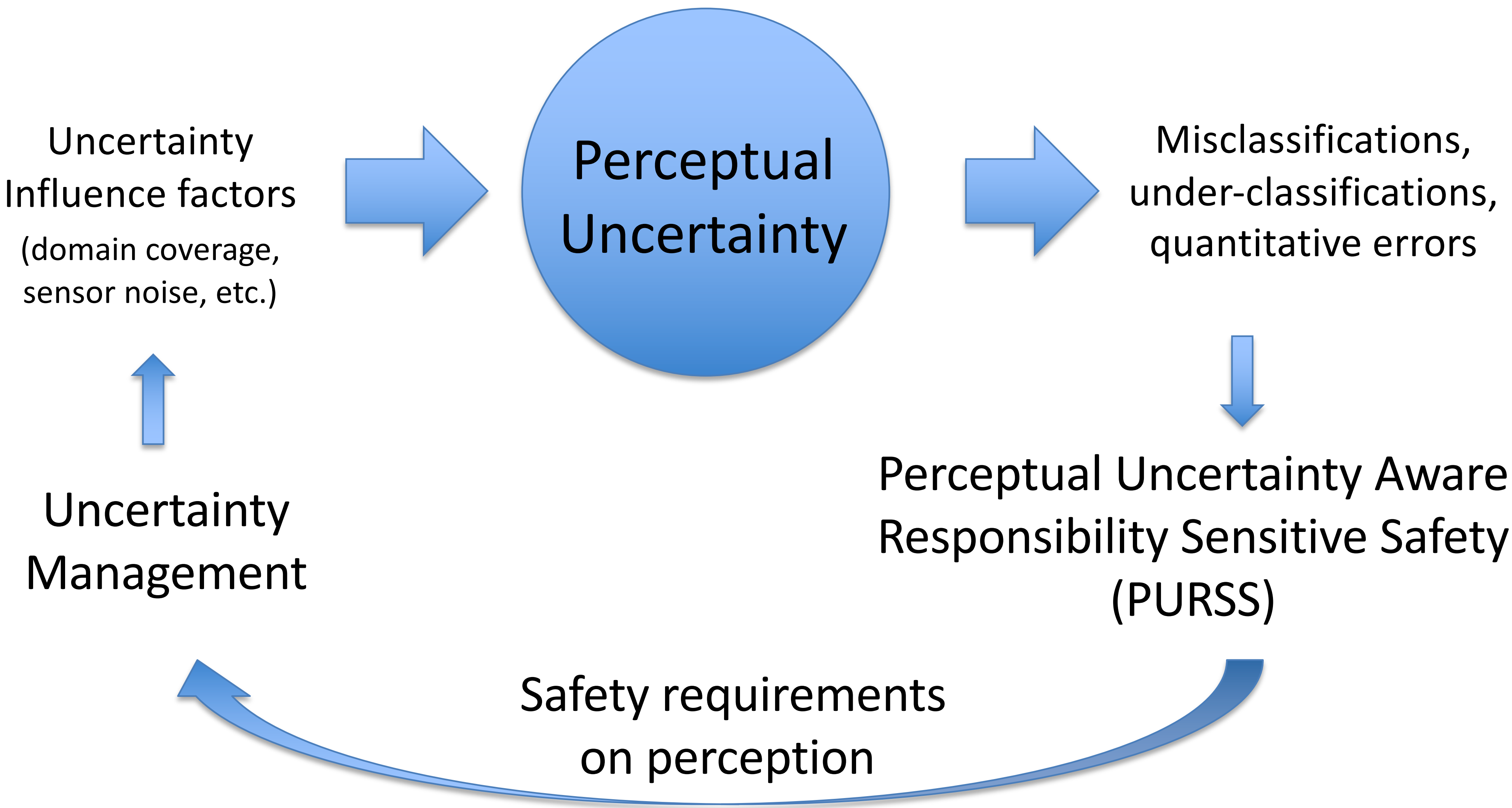


# Coming Soon: Canadian Adverse Driving Conditions Dataset





# Summary: Uncertainty-Centric Assurance of ML-Based Perception



# Insights and Challenges

- ML currently cannot be assured to certainty levels required for collision avoidance
  - ML is useful for longer-term, anticipatory risk reduction
- Perceptual uncertainty must be considered for the complete, fused perception and over time
  - E.g., different information becomes certain with different delays
- Out-of-distribution detection is still far from being useful in practice
- RSS leads to more conservative automated driving than human driving
  - E.g., negotiation in merging