Markus Stocker

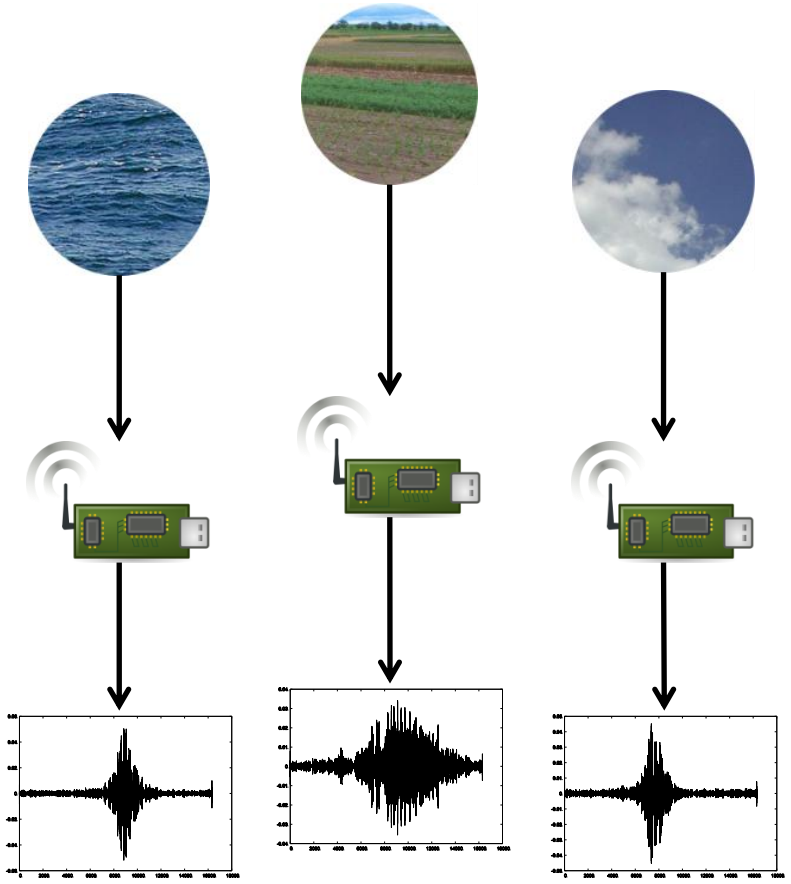
Environmental Informatics Seminar, Kuopio, March 13, 2012

Sensor networks: Too much data and not enough knowledge*

Approaching the problem using ontology

UNIVERSITY OF
EASTERN FINLAND

Environmental monitoring and sensor networks

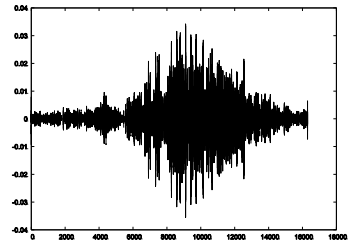


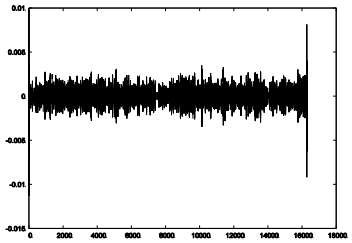
- Environmental monitoring
  - "Observe and record the conditions of the natural environment" [1]
  - Involves *measurement* as the "process of empirical, objective, assignment of numbers to the properties of objects and events of the real world in such a way as to describe them" [2]
  - Sensor is a device that performs measurement over time
- Sensor layer
  - Physical layer, hardware
  - (Heterogeneous) sensor network
- Data layer
  - Raw spatiotemporal measurement data
  - Potentially high frequency

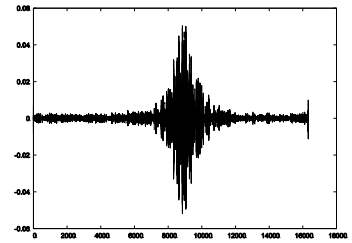
Too much data …

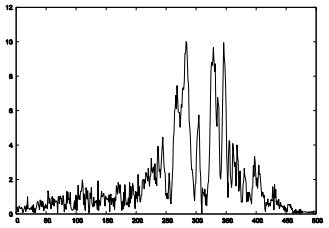
- Two examples for vibration and camera sensors
  - Vibration sensor at 2 kHz sampling frequency
  - Camera sensor at variable 1-3 fps
- One vibration sensor and one camera sensor
  - 710 hours of measurement, i.e. ~30 days
  - 5,004,130,000 vibration sensor measurement values
  - 1.7 TB data for vibration and camera sensor
- Three vibration sensors and one camera sensor
  - 6 hours of measurement
  - 128,888,596 vibration sensor measurement values
  - 25,076 camera sensor image files

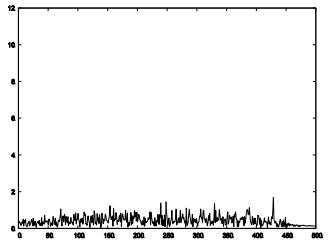
… and not enough knowledge

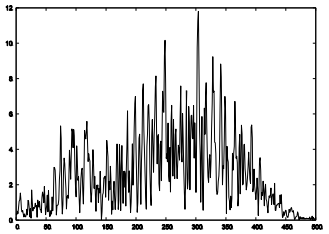










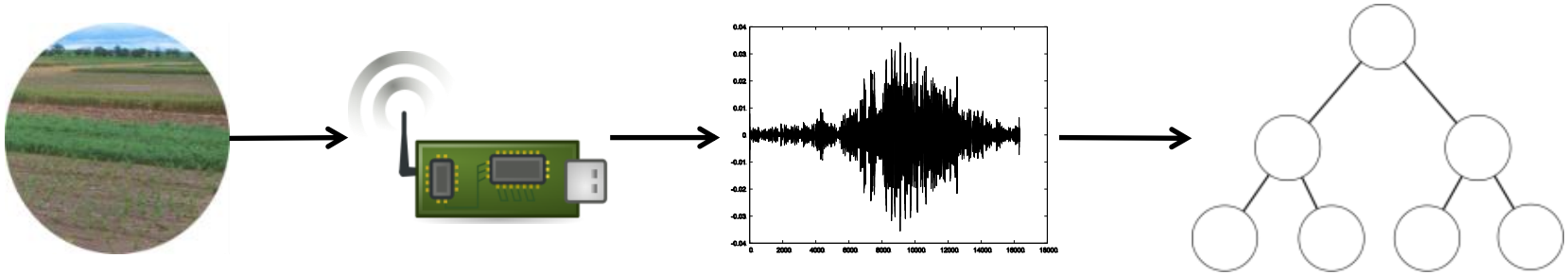


Challenges

- Data acquisition, processing, storage, and retrieval
  - Heterogeneous communication technology, protocols, data types, etc.
  - Data management and retrieval systems
- Big data may not be *feasible*
  - "[B]eing able **to manage and process** […] those data is going to be a huge challenge" (National Ecological Observatory Network [3])
- Big data may not be *useful*
  - "[B]eing able […] to **make sense of** those data is going to be a huge challenge" [3]
- Data-rich and knowledge-poor environment

To make sense of sensor data

How to **acquire** knowledge



How to **represent** knowledge

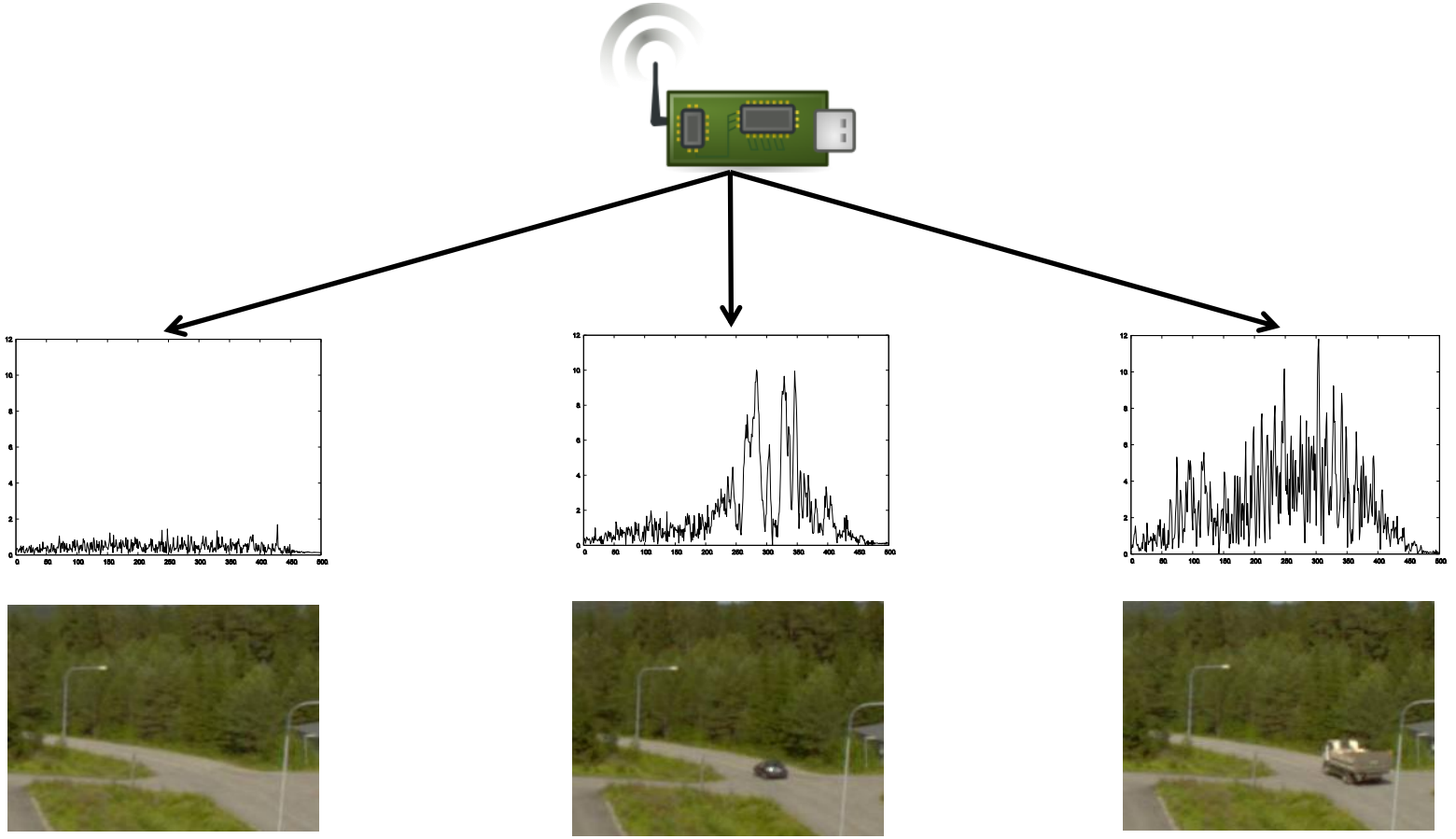
Ontology

- Defined as “an explicit specification of a conceptualization” [4]
- A means to formally represent knowledge of a domain
  - Concepts of some area of interest (i.e. domain)
  - Relations that hold among concepts
- Example (e.g. domain of road vehicles)
  - WeightCategory(light)
  - WeightCategory(heavy)
  - LightVehicle $\sqsubseteq$ Vehicle $\sqcap$ $\forall$hasWeightCategory.{light}
  - HeavyVehicle $\sqsubseteq$ Vehicle $\sqcap$ $\forall$hasWeightCategory.{heavy}
  - LightVehicle $\sqcap$ HeavyVehicle $\sqsubseteq$ $\bot$
  - VehicleObservation $\sqsubseteq$ Observation $\sqcap$ $\forall$featureOfInterest.Vehicle

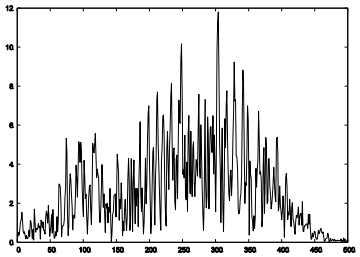
Monitoring road vehicles

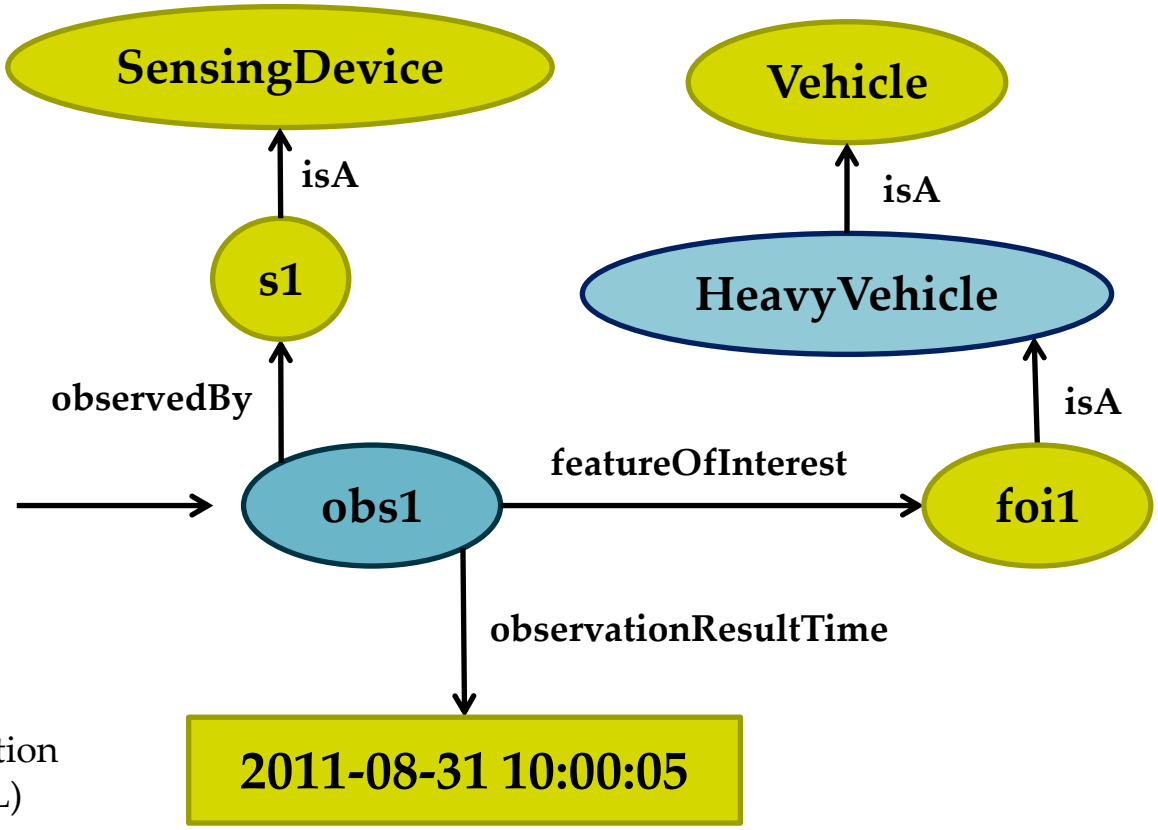
- Example used here
- Aim is detection and classification of vehicles
  - Is a vehicle observed by a sensor (now)?
  - If so, what vehicle type is it?
  - Machine learning (ML) classification
    - ~8 s of measurement used as pattern for classification
    - Supervised learning for pattern classification
- More aims
  - Infer when different sensors have observed the same vehicle
  - Infer vehicle velocity (speed, driving direction, driving side)

Patterns for classification (ML)



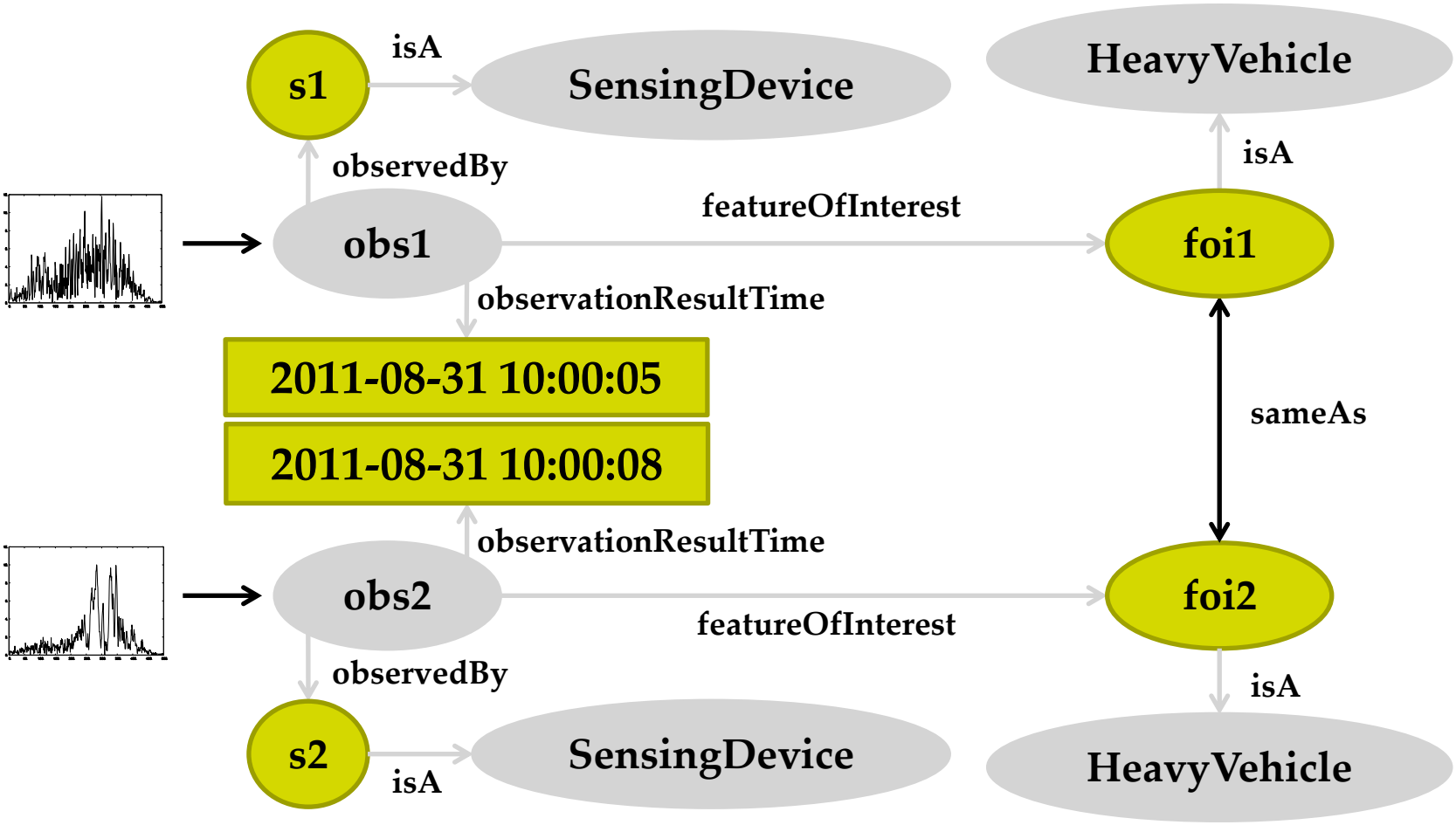
Knowledge representation
Observation for a vehicle



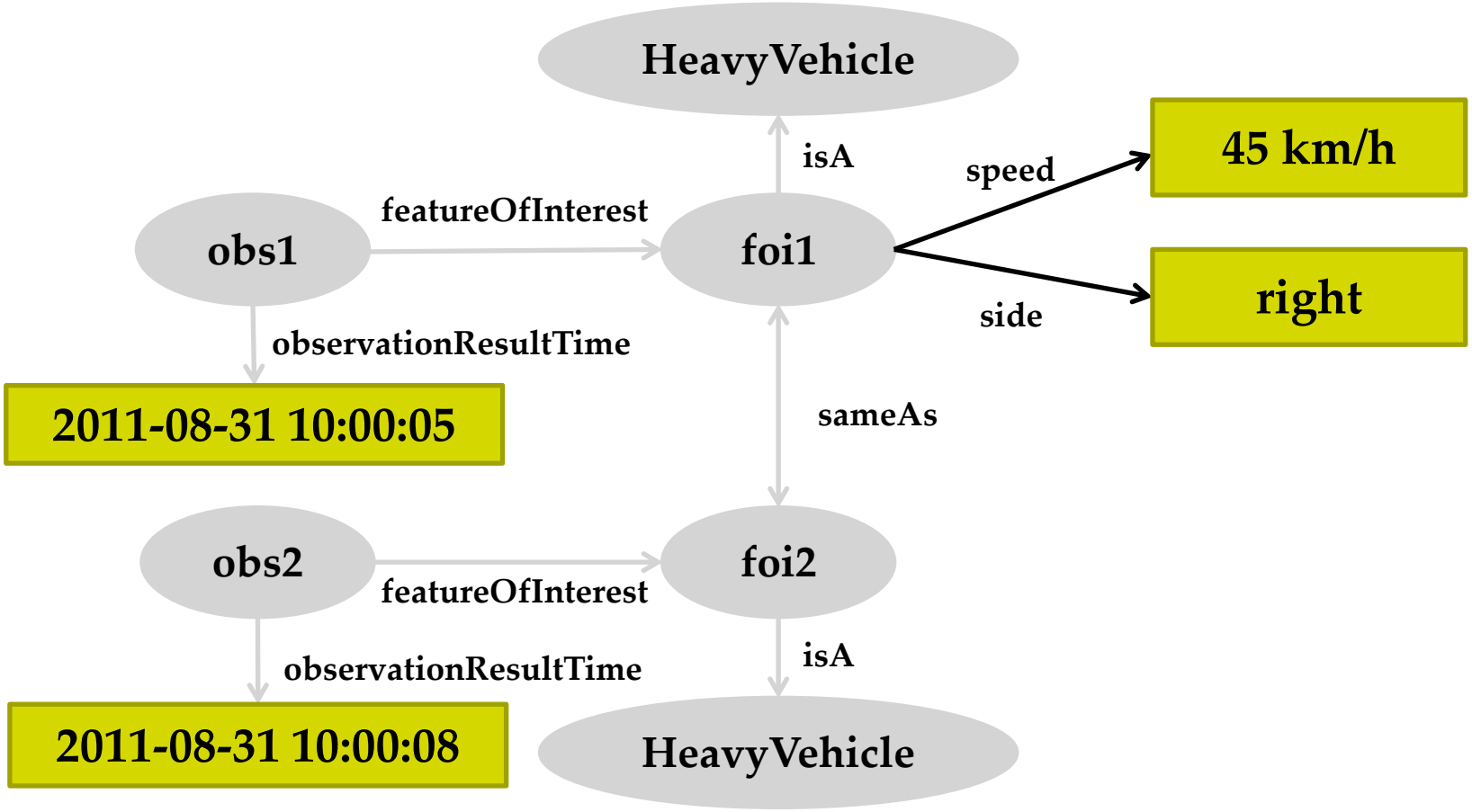


In blue: vehicle detection and classification (ML)

Knowledge representation
Two observations for same vehicle

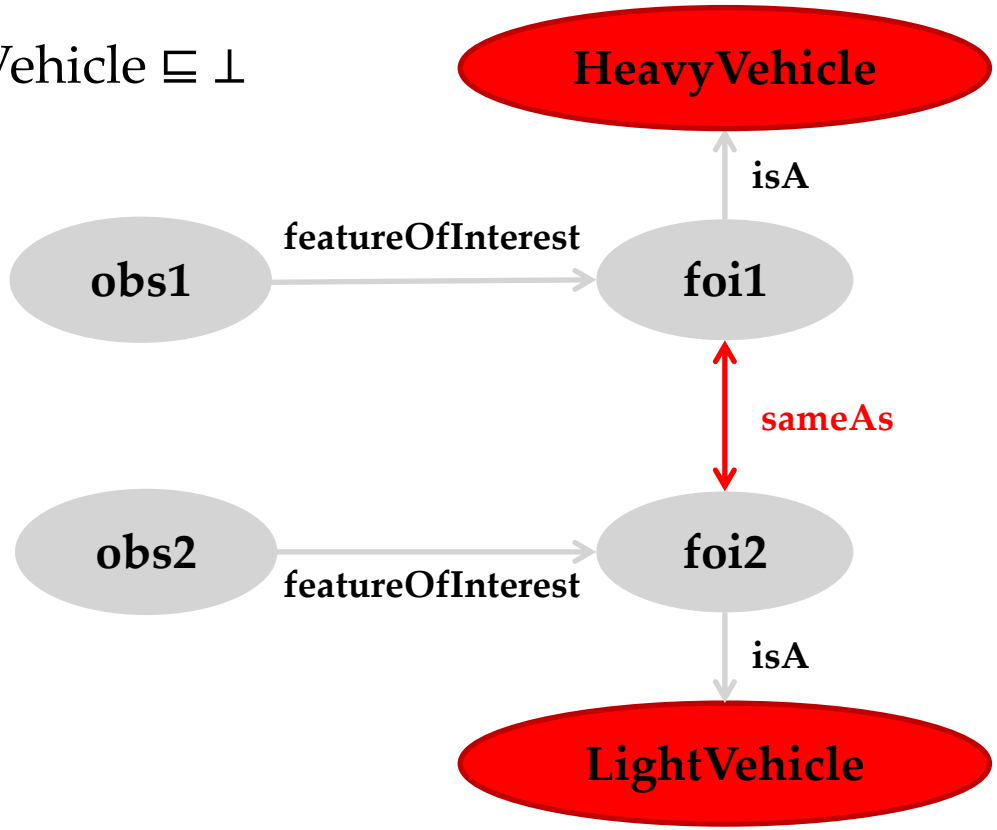


Knowledge representation
Inference of velocity



Knowledge representation
Consistency checking

$\text{LightVehicle} \sqcap \text{HeavyVehicle} \sqsubseteq \bot$



Conclusions

- Sensors and sensor networks can produce a lot of data
  - We are typically not interested in such data
  - We are interested in what those data tell
- We discussed automated workflow for
  - Sensor data acquisition and processing
  - Knowledge acquisition and representation
- Benefits of knowledge representation
  - Abstraction, inference, consistency checking, query, visualization
- Situation awareness? [5]
  - Perception: sensor network perceives status of elements in the environment
  - Comprehension: ontology as "holistic picture of the environment"
  - Projection: ontology reasoning
  - A methodology for situation assessment

References

(*) Sheth, A., C. Henson, and S. Sahoo (2008). Semantic Sensor Web. Internet Computing, IEEE, 12(4):78-83.

[1] Ioannis N. Athanasiadis and Pericles A. Mitkas. A methodology for developing environmental information systmes with software agents. In Advanced Agent-Based Environmental Management Systems. (Cortés, Ulises and Poch, Manel, Eds.), Birkhauser, 2009, pp. 119-138

[2] Finkelstein, L. Theory and Philosophy of Measurement. In Sydenham, P. and Thorn, R., Handbook of Measurement Science, Volume 1, Theoretical Fundamentals, pages 1-30. John Wiley & Sons, 1982.

[3] http://www.nature.com/news/2011/110809/full/476135a.html

[4] Thomas R. Gruber. A translation approach to portable ontology specifications. Knowledge Acquisition 5(2), 199-220 (1993)

[5] Mica R. Endsley. Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors, 1995, 37(1), 32-64