Towards Learning and Classifying Spatio-Temporal Activities in a Stream Processing Framework

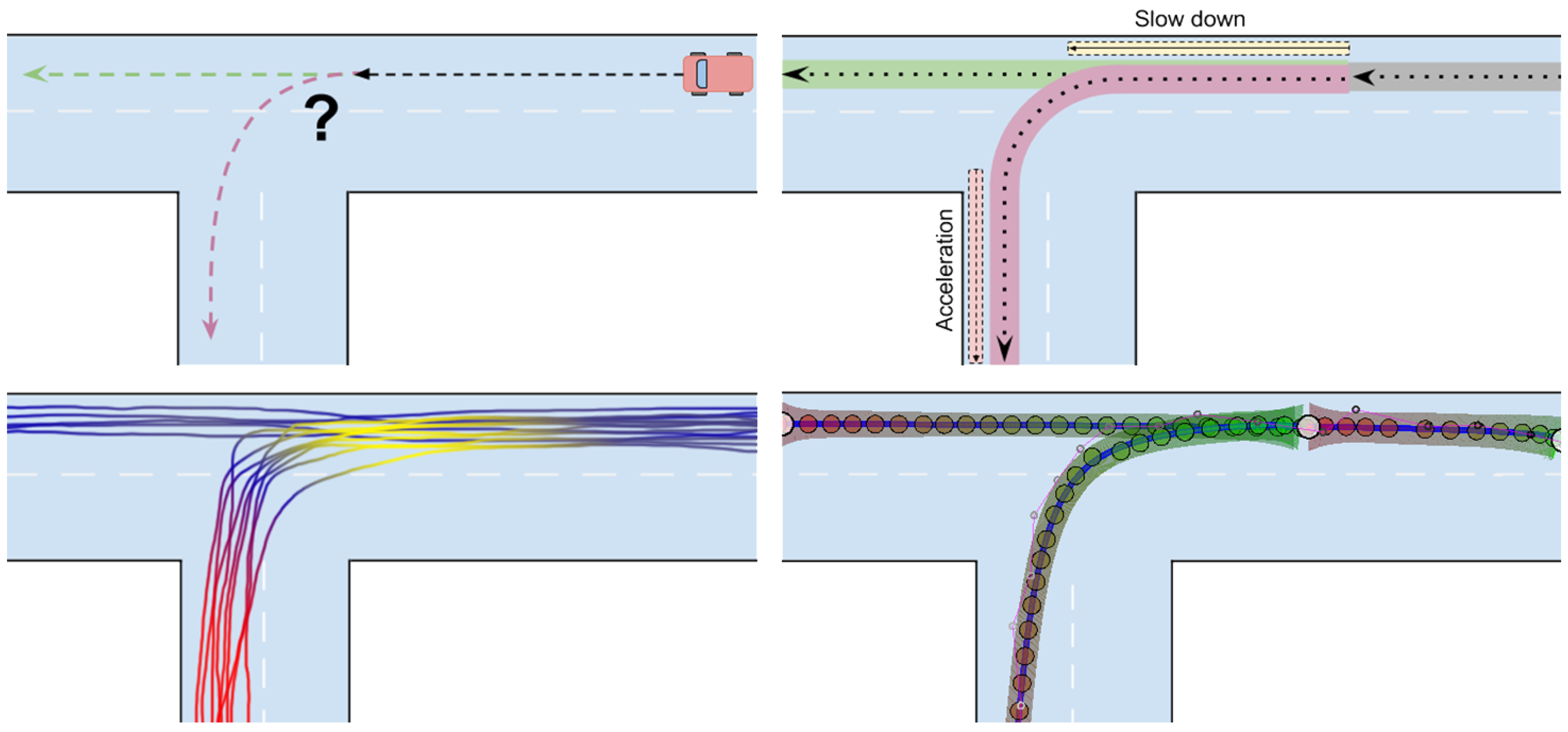
Mattias Tiger and Fredrik Heintz

Summary

Activity recognition is important for many applications and essential for situation awareness in real world environments. An important problem is unsupervised learning of activity representations and their context, allowing a system to quickly adapt to new phenomenon. To solve this problem we present a framework that can unsupervised learn discriminative activity models and their relations from observed trajectories in a state space. The activities are modeled using Gaussian processes and they are related in a Bayesian graph.

Problem Overview

Consider a T-crossing and the task of predicting if an observed car will turn or not. Changes in the behavior of the car are likely to occur before the turn, such as the car slowing down. At a certain point this difference in behavior is sufficient to separate them. The expected trajectory of cars will at this point split into two trajectories that should be possible to separate.



The figures above (left to right, top to bottom) show the scenario, expected dynamics, expected measurements and what the framework produces.

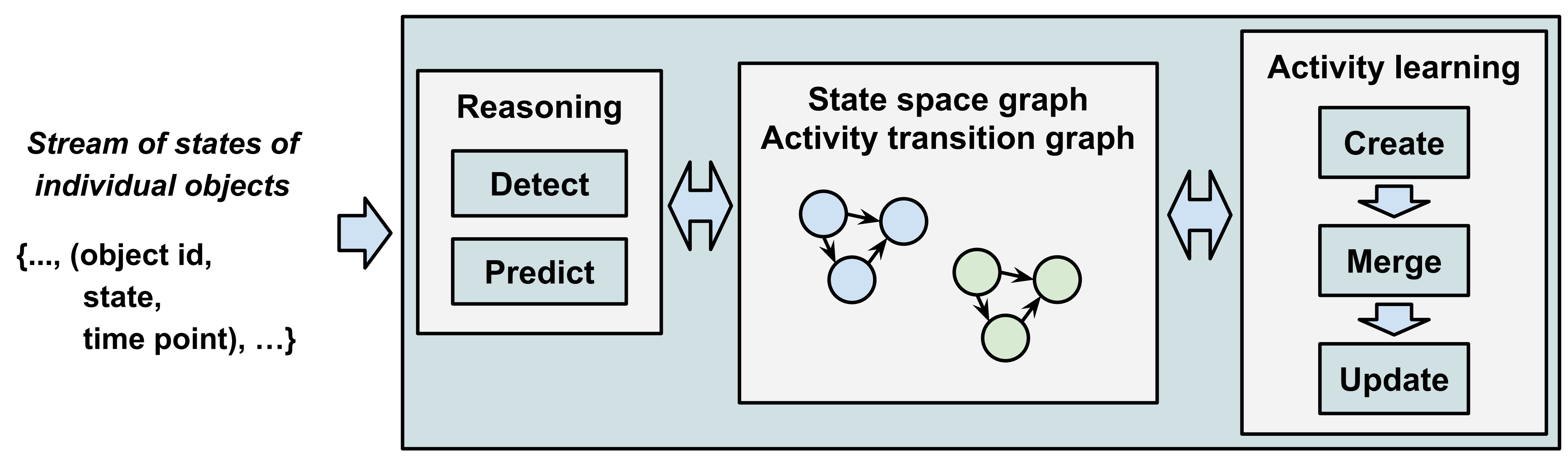
Activities

An *activity* can informally be defined as something a particular object does. In our case, an instance of an activity is represented by a state trajectory where some state variables are expected to change in a certain way while others may change freely. For example, an activity could be a slowdown where the velocity state variable is expected to decrease.

To model activities as continuous trajectories we use sparse Gaussian processes, which are probability distributions over functions. The activities are atomic and connected to each other at branching points, forming a state space graph. The transition probabilities between different activities form an activity transition graph. Both are modeled as causal discrete Markov chains.

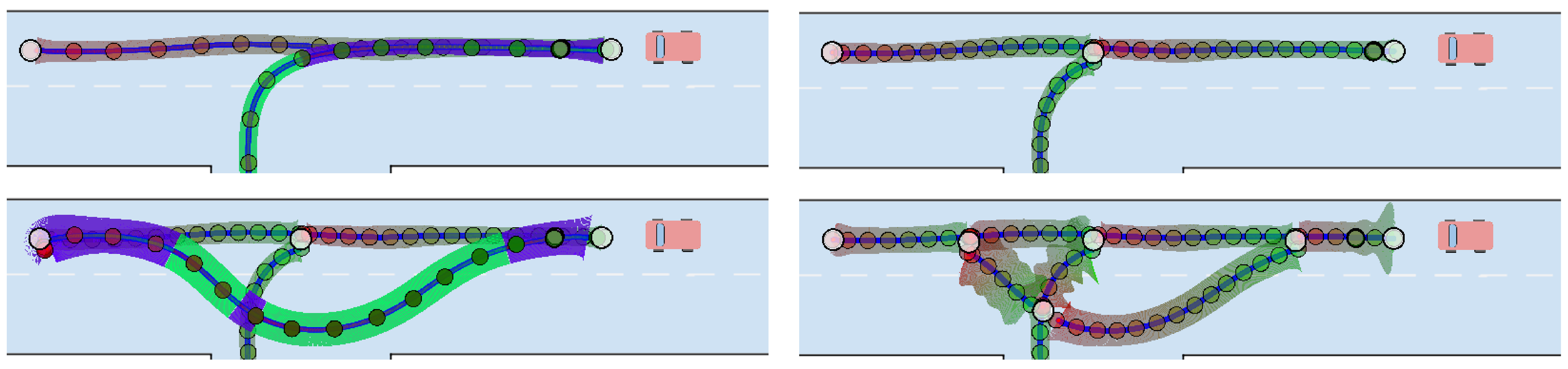
Learning and Classification Framework

The proposed framework learns activities and their relations based on streams of temporally ordered time-stamped probabilistic states for individual objects in an unsupervised manner, utilizing stream processing capabilities provided by frameworks such as DyKnow[1]. Using the learned activities, the framework classifies the current trajectory of an object to belong to the most likely chain of activities. The framework can also predict how abnormal the current trajectory is and how likely future activities are.



Learning and Classification

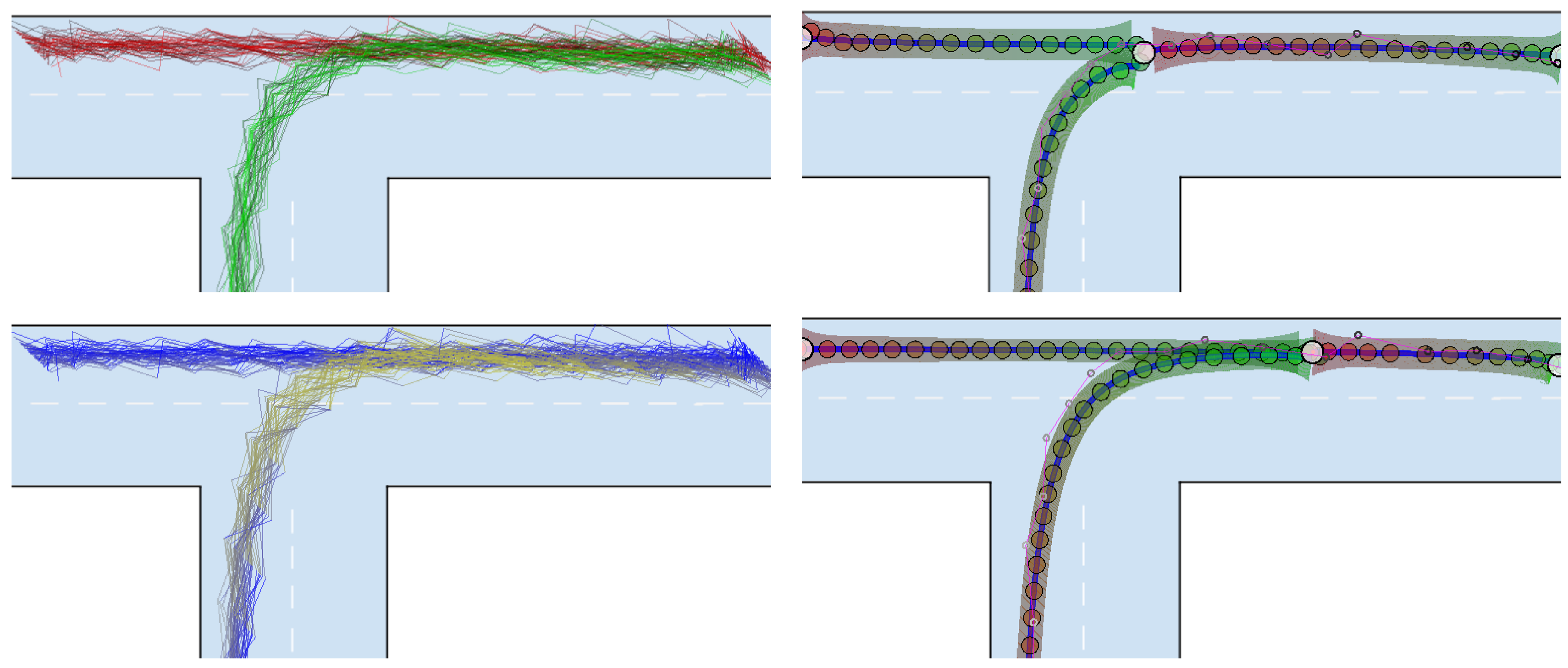
To learn activities from observed trajectories, the framework segments trajectories into sequences of activities, where each activity differs from all other activities currently known. Activities are segmented such that no two activities can have overlapping intervals in the state space, where the margin for overlap is controlled by a user set detail threshold. Comparison of activities is performed by calculating the local likelihood[2] at sufficiently many points.



The figures above (left to right, top to bottom) illustrate the segmentation of activities from 2D position observations of different behaviors in the order of driving straight, left turn and overtake. Blue indicates intervals explained by an activity, green indicates not explained.

Preliminary Results

To verify that our approach works as expected, we have done a case study with a T-crossing scenario. The simulated data set contains two types of trajectories, straight trajectories and left turning trajectories, 50 of each. These trajectories are provided to the framework in a random order. With position and constant velocity the turning activity breaks up a bit earlier than with only position, and even more so with a slow down taking place. The mean processing time per observed trajectory is for both data sets less than 0.55 ± 0.08 seconds. A single core on a 2.5GHz i5-4300U processor was used.



The figures above show at the top constant speed, at the bottom a slowdown. Data sets to the left, converged models to the right. Blue indicates normal speed and yellow indicates lower speeds.

Conclusions

The framework is able to unsupervised learn connected atomic activity representations from observed trajectories by discriminating between intervals that are and aren't similar enough to previously learned activities.

The framework cannot yet learn complex activities as discrete entities other than indirectly by parts, i.e. by learning chains of primitive activities. Neither are contexts such as a T-crossing learned, instead configurations of different activities are learned for each T-crossing instance. The current framework is designed to provide the building blocks for these extensions.

References

[1] Fredrik Heintz, Jonas Kvarnström, and Patrick Doherty. Bridging the sense-reasoning gap: DyKnow – stream-based middleware for knowledge processing. *Advanced Engineering Informatics*, 24(1):14–26, 2010.

[2] K. Kim, D. Lee, and I. Essa. Gaussian process regression flow for analysis of motion trajectories. In *Proc. ICCV*, 2011.

This work is supported by grants from the Swedish Foundation for Strategic Research (SSF) project CUAS, the Swedish Research Council (VR) Linnaeus Center CADICS, the ELLIIT Excellence Center at Linköping-Lund for Information Technology, and the Center for Industrial Information Technology CENIIT.