Scenario 2 Vector: Scenario Description Language Based Embeddings for Traffic Situations

Aron Harder(ah2ph), Jaspreet Ranjit (jr4fs), Manisha Sudhir (ms8jd)

Problem Definition

Introduction

1. Objective:

- a. We propose **Scenario2Vector** a Scenario Description Language (SDL) based embedding for traffic situations that allows us to automatically search for similar traffic situations from the large data-sets obtained from different AVs
- b. *Input:* traffic scenario (i.e lane change)
- c. **Output:** A similar traffic scenario (semantically and temporally) in a different dataset, standard metric of comparison

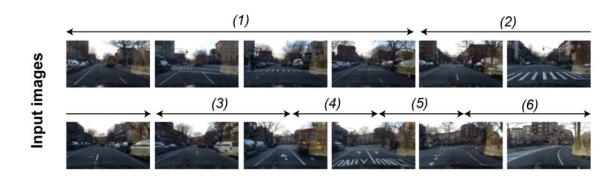
2. Key questions:

a. Given a traffic scenario (5-15 second video clip), can we query another database for a similar video using spatio temporal properties of that video, and a high level video description as the label?





Berkeley Deep Drive-X Dataset



Action description:

- (1) The car is driving
- (2) The car is moving into the right lane
- (3) The car moves back into the left lane
- (4) The car drives in the left lane
- (5) The car moves into the right lane
- (6) The car proceeds down the road

Action justification:

as there is nothing to impede it.

because it is safe to do so.

because the school bus in front of it is stopping.

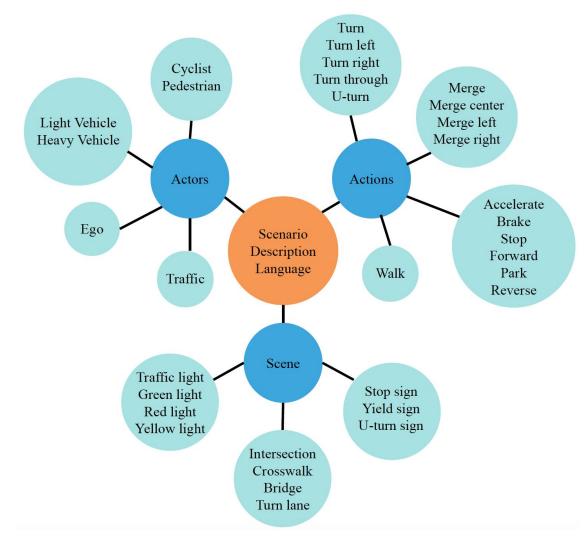
in order to pass the school bus.

since it has now passed the school bus and it is taking the right fork.

as there is nothing in its way.

Methodology

Scenario Description Language (SDL)



SDL Extraction Process



(1) BDDX Ground Truth: The car is stopped at a crosswalk as the light is red and pedestrians are crossing the street



(2) Actions and Scene Elements



 $S = \langle AT, AC, SC \rangle$ AT = [Ego, Pedestrian]

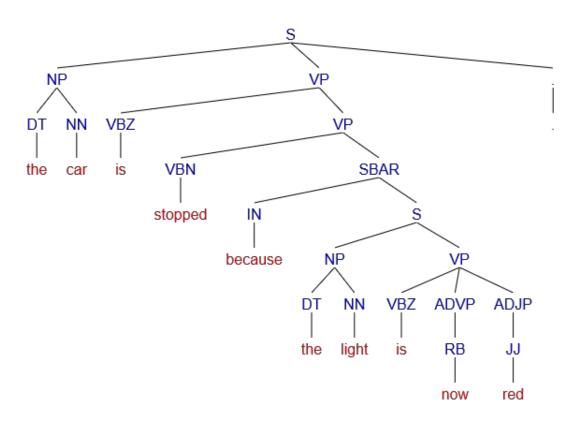
AA = [Stop Walk] SC = [Red light, Crosswalk]

[Ego Pedestrian]

(3) SDL Extraction (4

(4a) SDL Matrix Embedding

SDL Extraction Details



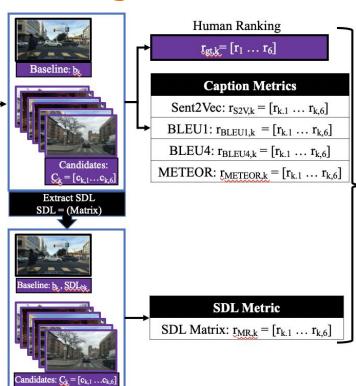
Experiments

Similarity Ranking

Traffic Scenario Similarity Dataset

- 100 scenario samples – subset of BDDX dataset
- Each scenario sample contains a baseline and 6 candidates
- Candidates ranked by humans based on similarity to baseline

 $SDL_{Ck} = [SDLc_{k,1}...SDLc_{k,6}]$



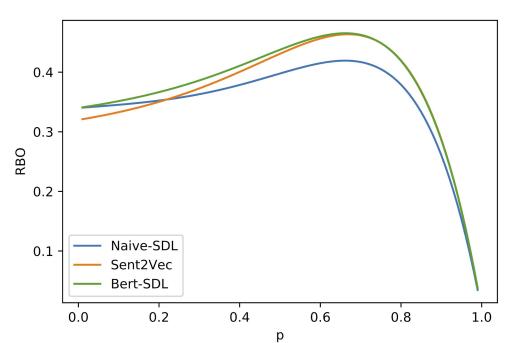
	Results	
Metric	Rank Biased Overlap	P-value
SDL Matrix (Bert)	0.46534	0.65
Sent2Vec	0.46356	0.66
SDL Matrix	0.41932	0.65

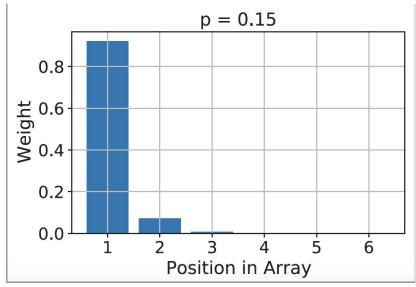
Rank

Biased

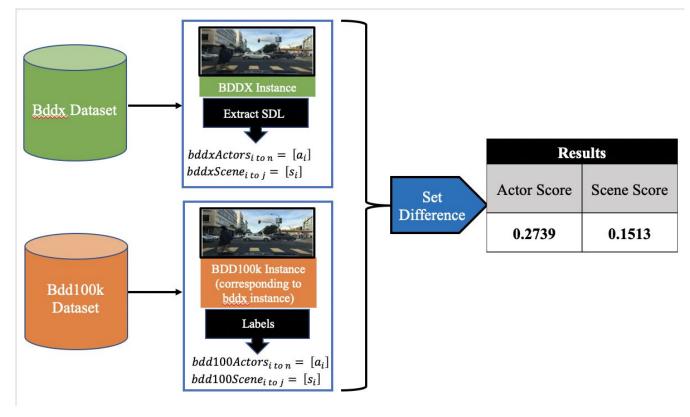
Overlap

Similarity Ranking Results





Object Detection vs. SDL Extraction Setup and Results



Object Detection vs. SDL Extraction Analysis

BDD100k

Actors

• Train

Scene

- Tunnel
- Residential
- Parking lot
- City Street
- Gas stations
- Highway

Actor

- Light Vehicle
- Heavy Vehicle
- Cyclist
- Pedestrian

Scene

- Traffic light
- Traffic sign
- Crosswalks

Scenario Description Language

Actor

- Ego
- Traffic

Scene

- Intersection
- Crosswalk
- Bridge
- Stop sign
- Yield sign
- U-turn
- Turn lane