Online Decision-Making for Scalable Autonomous Systems

Kyle H. Wray\textsuperscript{1,2}  Stefan J. Witwicki\textsuperscript{2}  Shlomo Zilberstein\textsuperscript{1}

\textsuperscript{1}University of Massachusetts Amherst
\textsuperscript{2}Nissan Research Center - Silicon Valley

August 23, 2017
Motivation: Autonomy in the Open World

Autonomous Vehicle (AV) Intersection Decision problem

Complex interactions with multiple, simultaneously encountered entities
Motivation: Autonomy in the Open World

Autonomous Vehicle (AV) Intersection Decision problem

Pedestrian #1

Vehicle #2

Vehicle #1

AV

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
Motivation: Why Is It Hard?

- Right of way requires careful monitoring of arrival times of other vehicles, can be ambiguous
- A vehicle that has the right of way could be blocked (e.g., by pedestrians or bicyclists)
- Entities can convey beliefs and intention via actions (e.g., edging) or signals (e.g., blinkering)
- The number of entities that may interact with an AV at an intersection could be very large
- Factoring all the possible entity configurations in every possible intersection is intractable
AV Intersection Decision-making: Problem Statement

- **Problem:** Long-term urban deployment of AVs requires the critical *mid-level decision-making*
  - second-to-second stop, yield, edge, or go decisions

- **Overall Goal:** Produce a *general solution* to intersection decision-making, for both vehicles and pedestrians
  - n-way stop signs, traffic lights, T-intersections, uncontrolled left turns, right turn on red, crosswalkers, jaywalkers, etc.

- **AV Objective:** Reach other side of intersection without causing conflicts (physical or social)
  - respect traffic laws and social norms
AV Intersection Decision-making: Formulated as POMDPs

- **Sequential Decision Problem:**
  - dynamically changing state of each entity (vehicle, pedestrian)
  - "blocking" and "priority" relations change over time

- **Multiple Sources of Uncertainty:**
  - behavior of other vehicles: *stop*, *yield*, *go*, etc.
  - noisy sensors: detection, movement, lane, blocking, etc.
  - occlusions: possibly hidden vehicles or pedestrians

- **Key Challenges:**
  - unbounded and changing number of relevant entities
  - entities unknown a priori, require real time response
  - extremely large state space, scalability and complexity
MODIA: High Level View

Multiple Online Decision-components with Interacting Actions

- Describe each potential pairwise interaction problem using a generic Decision-Problem (DP)
- Model each DP using a POMDP and solve it offline
- Instantiate DPs for each encountered entity and create an active Decision-Components (DC)
- Aggregate DC action recommendations into a final action using an Executor

Objective: Minimize executor final action regret among DCs
MODIA: Decision-Problem (DP)

- **DP = specific abstract problem**
  - e.g., “intersection decisions with one other vehicle approaching from the right”

- **DPs are known a priori**
  - typically solved offline (e.g., < 30 DPs, each < 1 min to solve)

- **DPs are currently solved as POMDPs**
  - architecture is modular; other models fit naturally

**Definition:** $\langle P, A \rangle$
- $P = \{P_1, \ldots, P_k\}$ is a set of $k$ decision-problems
- $A$ are the *actual* final actions taken by the agent
MODIA: AV DPs (Example) (1/3)

State factors and actions

- **States:** \( S = S_{av}^L \times S_{av}^T \times S_{ov}^L \times S_{ov}^T \times S_{ov}^b \times S_{ov}^p \)
  - AV position \( S_{av}^L = \{ \text{Approaching, At, Edged, Inside, Goal} \} \)
  - AV wait time \( S_{av}^T = \{ \text{Short, Long} \} \)
  - other position \( S_{ov}^L = \{ \text{Empty, Approaching, At, Edged, Inside} \} \)
  - other wait time \( S_{ov}^T = \{ \text{Short, Long} \} \)
  - blocking (do trajectories intersect?) \( S_{ov}^b = \{ \text{Yes, No} \} \)
  - priority (who arrived first?) \( S_{ov}^p = \{ \text{AV, OtherVehicle} \} \)

- **Actions:** \( A = \{ \text{Stop, Edge, Go} \} \)
  - note: the “edge” action means move slowly forward.
MODIA: AV DPs (Example) (2/3)

Illustration of some state factors with two instantiated DCs

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
MODIA: AV DPs (Example) (3/3)

Illustration of two DPs for AVs
MODIA: Decision-Component (DC)

- **DC = instantiate and execute a DP**
  - many copies of an intersection DP could each form a DC

- **DC are unknown a priori**
  - created online while driving, based on world

- **DCs are copies of POMDP DPs with internal beliefs**
  - \( DC = \langle DP, policy, current \text{ belief} \rangle \)

- **Definition:** \( \langle C, \phi, \tau \rangle \)
  - \( C = \{ C_1, \ldots, C_n \} \) is a set of \( n \) decision-components
  - \( \phi: C \to \mathcal{P} \) map DCs to their originating DP
  - \( \tau: C \to \mathbb{N}^2 \) such that \( \tau(C_i) = [\tau_s(C_i), \tau_e(C_i)] \) (start/end times)

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein
Online Decision-Making for Scalable Autonomous Systems
MODIA: AV DCs (Example)

Three instantiated DCs: two vehicles and one pedestrian

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
MODIA: Executor

- **Executor** = aggregate DC action recommendations
  - currently, lexicographic preference for safest action ("LEAF")

- **Defined offline, aggregated online, known a priori**
  - definition allows for *online learning* of the executor

- **Definition:**  $\epsilon : \tilde{A} \rightarrow \mathcal{A}$ with $\tilde{A} = (\bigcup_i A_i)^*$
  - DCs are unknown a priori
  - executor *must* be capable of aggregating *any* action tuple
MODIA: AV Executor (Example)
MODIA: Objective

- **Objective** = minimize regret among DCs
  - e.g., many DCs want 'go', but one says 'stop' \(\Rightarrow\) 'stop' taken

- **Executor selection induces particular regrets**
  - allows us to prove that particular executors will minimize regret in particular MODIA

- **Definition**: total regret \(R^h_\epsilon\); one-step regret \(r^t_\epsilon\):

\[
r^h_\epsilon = \sum_{i \in I^t} Q_i(b^{t_i}_i, \pi_i(b^{t_i}_i)) - Q_i(b^{t_i}_i, \epsilon_i^{-1}(a^t))
\]

with \(R^h_\epsilon = \sum_{t=1}^h r^t_\epsilon\)

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
MODIA: Complete Picture (Example)

Motivation

Problem Statement

MODIA

Analysis

Experiments

Conclusion

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
Analysis: Properties of MODIA

- **Proposition 1:** If all instantiations are known, then computing the optimal solution has *polynomial* complexity (i.e., MODIA is tractable)

- **Proposition 2:** If only an instantiation model is known, then computing the optimal solution has *exponential* complexity (i.e., monolithic POMDP is intractable)

- **Proposition 3:** For a class of MODIA (risk-sensitive), a specific simple executor (LEAF) is the optimal solution (i.e., lexicographic action selection minimizes regret)
# MODIA: Experimental Results

Experiments using a high-fidelity industrial simulator

| Name            | RS | V | P | PI | $|C|$ | $\mathcal{M}$ | $I$ | $N$ |
|-----------------|----|---|---|----|----|-------------|-----|-----|
| Crosswalk Pedestrian | 4  | 0 | 1 | 1  | 4  | 21.1        | 16.7| 30.1|
| Vehicle & Pedestrian    | 3  | 1 | 1 | 1  | 3  | 16.8        | 13.6| 37.1|
| Walk & Run Pedestrians  | 3  | 1 | 2 | 2  | 6  | 19.1        | 13.3| 23.3|
| Multi-Vehicle Interaction| 4  | 2 | 0 | 2  | 5  | 19.0        | 13.2| 20.9|
| Bike Crossing         | 3  | 0 | 1 | 1  | 3  | 16.4        | 13.8| 19.8|
| Jay Walker           | 4  | 0 | 1 | 1  | 4  | 17.7        | 14.4| 24.3|

Results for six intersection problems described by the number of road segments (RS), vehicles (V), pedestrians (P), and potential incidents (PI). MODIA AV $\mathcal{M}$ (number of DCs $|C|$) is compared with two baselines, ignorant $I$ and naive $N$, using their intersection completion times (seconds).
MODIA: Experimental Results

Initial testing on Nissan’s experimental autonomous vehicle

Kyle Wray, Stefan Witwicki, Shlomo Zilberstein

Online Decision-Making for Scalable Autonomous Systems
Conclusion

• Present a general scalable decision-making model: MODIA
• Analyze MODIA’s complexity and optimality
• Introduce an efficient optimal solution (LEAF) for a class of MODIA (risk-sensitive)
• Describe an example solution for the challenging domain of AV intersection decision-making
• Demonstrate successful use of MODIA on a fully-functional AV prototype
Questions?

- For additional Information: Resource-Bounded Reasoning Lab University of Massachusetts Amherst
  http://rbr.cs.umass.edu

- We thank our sponsors: NSF and Nissan Motor Corporation