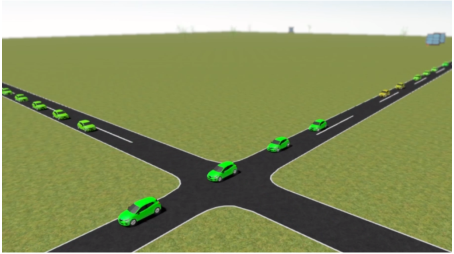
Scalable Coordination of Autonomous Vehicles in Networks

Jeroen van Riel

January 2025

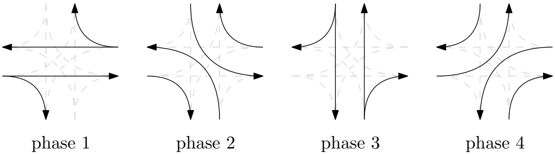
Autonomous vehicles



- Coordination of autonomous vehicles
  - Better guarantees on safety
  - Potentially reduce economic costs
  - No need for traffic lights

Related literature

- Traffic light control with deep reinforcement learning
  - Policy to control *phase* of signal

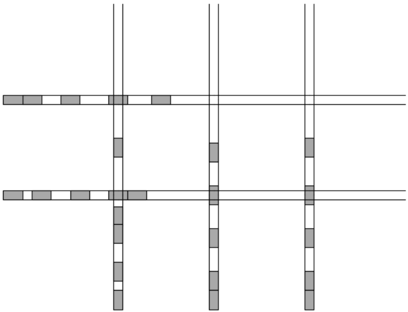


  - Based on microsimulation (e.g., SUMO, VISSIM)
  - Multi-agent perspective: each intersection has some degree of autonomy

Related literature

- No traffic lights (*autonomous intersections*)
  - Autonomous (rather "automated") vehicles
  - Coordination for
    - Collision avoidance
    - Efficiency
  - Locus of control: central $\leftrightarrow$ distributed
  - Central control $\implies$ optimal control problem

Networks of autonomous intersections



- Fixed routes
- Double integrator vehicle dynamics
- Collision-free trajectories

Problem formulation

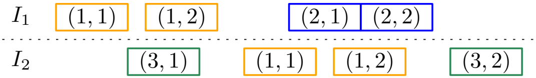
- Algorithmic challenges
  1. Safety with respect to collisions
  2. Scalability to large urban networks
  3. Learn from interaction with the system
- Research questions
  1. Formulate as job-shop scheduling
  2. Apply Deep Reinforcement Learning (DRL)

Research approach

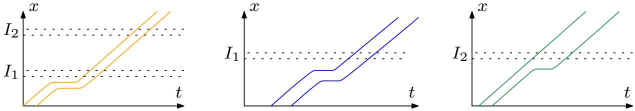
- Decomposition
- Vehicle scheduling problem
- DRL heuristic

Decomposition

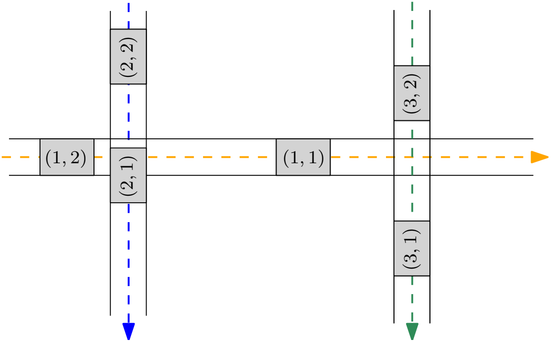
- Upper-level vehicle scheduling problem
  - Mixed-Integer Linear Programming (MILP)



- Lower-level trajectory optimization problem
  - Direct transcription $\rightarrow$ linear programming

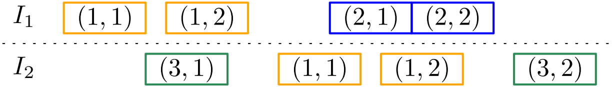


Vehicle scheduling problem

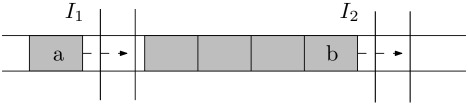


- Order vehicles at intersections
- Intersection occupied by at most one vehicle

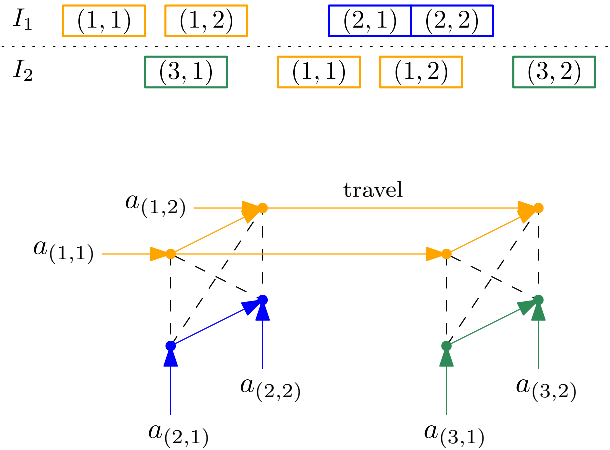
Vehicle scheduling problem



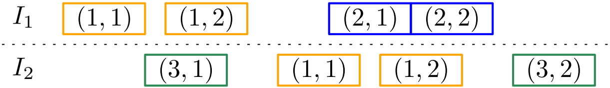
- In which order should vehicles cross intersections?
- Travel constraints model minimum travel time
- Buffer constraints to prevent overcrowding at lanes

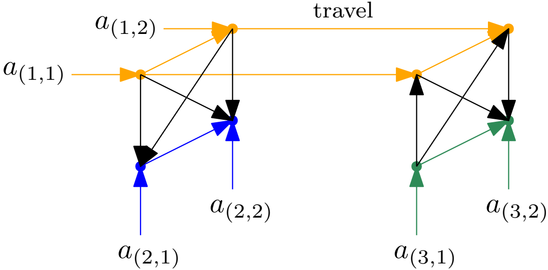


Vehicle scheduling problem

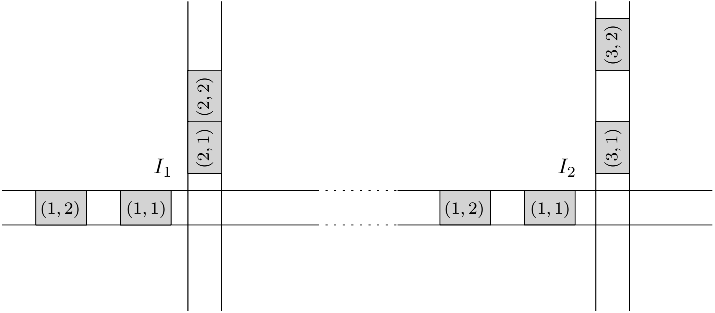


Vehicle scheduling problem

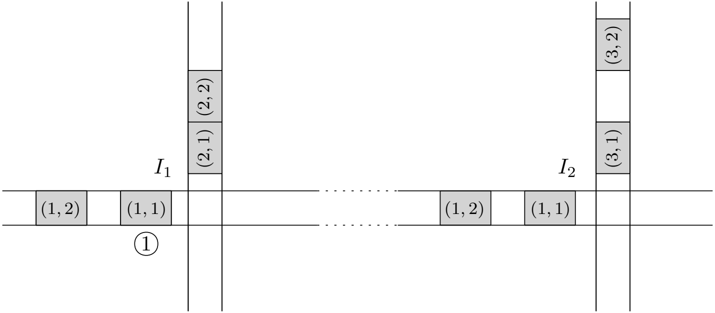




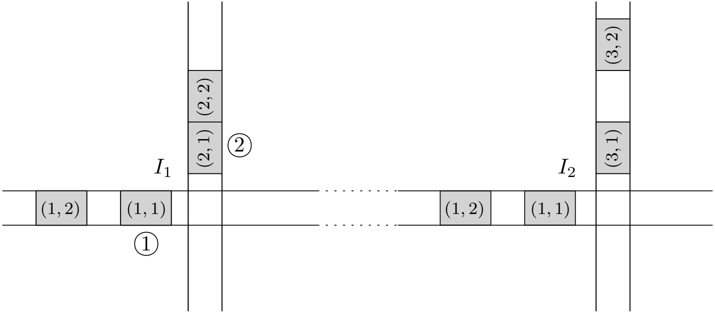
DRL heuristic



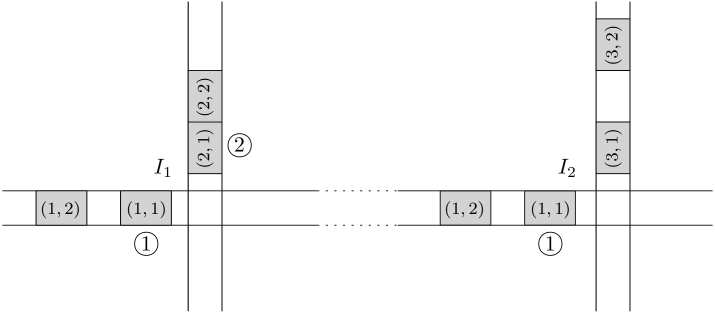
DRL heuristic



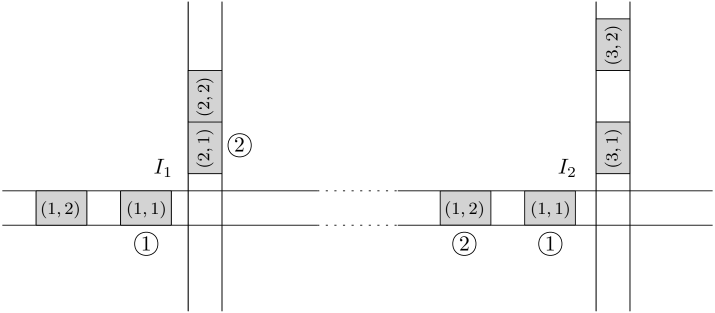
DRL heuristic



DRL heuristic



DRL heuristic



DRL heuristic

- Step-by-step construction of crossing order
- Partial solutions encoded as disjunctive graph augmented with lower bounds on crossing times
- Parameterize ordering policy based on graph neural network embedding of augmented disjunctive graph

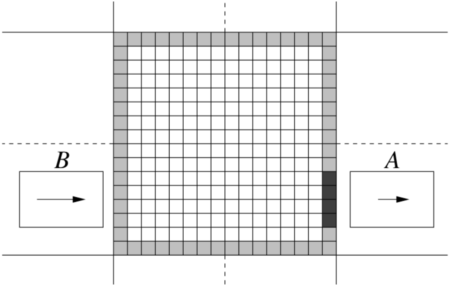
- Imitation learning from MILP solutions
- Reinforcement learning from vehicle delay reward signal

Appendix: Related literature

- Autonomous intersections
- Neural combinatorial optimization

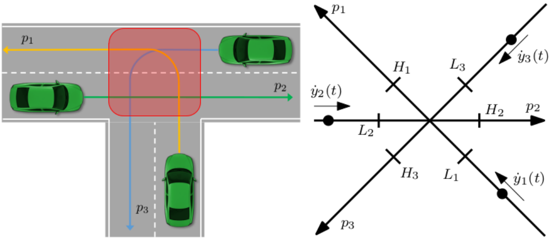
Autonomous intersections

- "Autonomous Intersection Control" (Dresner & Stone)
  - Single intersection
  - Time slot reservation-based protocol
  - Central intersection manager



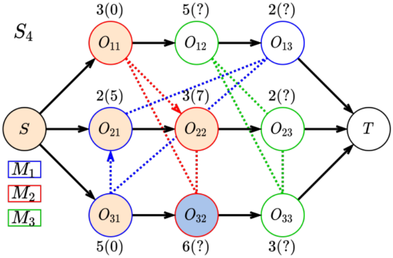
Autonomous intersections

- "Approximate Optimal Coordination" (Hult et al.)
  - Single intersection
  - Single vehicle per lane
  - Explicit collision-avoidance constraints

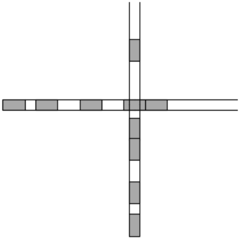


Neural combinatorial optimization

- "Learn to dispatch" (Zhang et al.)
  - Job-shop scheduling problem
  - Dispatch next operation
  - Policy using Graph Isomorphism Network (GIN)



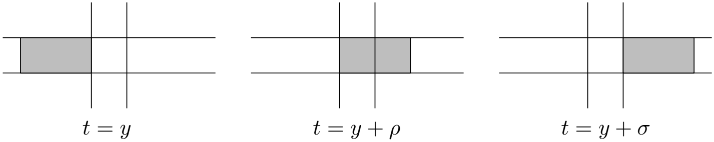
Appendix: Single intersection



- Notation
- Upper-level crossing time scheduling
- Lower bound on starting times
- Imitation learning with neural policy
- Lower-level trajectory optimization

Notation

- vehicle indices $\mathcal{N}$
- $y(i)$ is crossing time of vehicle i
- r_i earliest crossing time of vehicle i



- i and j same lane: $y(i) + \rho \leq y(j)$
- i and j distinct lanes: $y(i) + \sigma \leq y(j)$ or $y(j) + \sigma \leq y(i)$

Upper-level crossing time scheduling

- conjunctive constraints $\mathcal{C}$
- disjunctive (conflict) constraints $\mathcal{D}$

$$
\begin{aligned}
\min_{y} \quad & \sum_{i \in \mathcal{N}} y(i) & \\
\text{s.t.} \quad & r_i \leq y(i), & \text{for all } i \in \mathcal{N}, \\
& y(i) + \rho \leq y(j), & \text{for all } (i,j) \in \mathcal{C}, \\
& y(i) + \sigma \leq y(j) \text{ or } y(j) + \sigma \leq y(i), & \text{for all } (i,j) \in \mathcal{D}
\end{aligned}
$$

Upper-level crossing time scheduling

- Formulate as mixed-integer linear program (MILP)
- Introduce binary decision variables γ_{ij}
- Use big-M technique

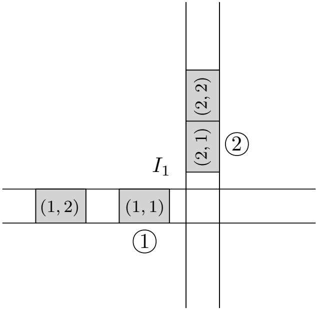
$$
\begin{aligned}
\min_{y} \quad & \sum_{i \in \mathcal{N}} y_i \\
\text{s.t.} \quad & r_i \leq y_i, && \text{for all } i \in \mathcal{N}, \\
& y_i + \rho_i \leq y_j, && \text{for all } (i,j) \in \mathcal{C}, \\
& y_i + \sigma_i \leq y_j + \gamma_{ij} M, && \text{for all } (i,j) \in \mathcal{D}, \\
& y_j + \sigma_j \leq y_i + (1 - \gamma_{ij}) M, && \text{for all } (i,j) \in \mathcal{D}, \\
& \gamma_{ij} \in \{0,1\}, && \text{for all } (i,j) \in \mathcal{D}
\end{aligned}
$$

Lower bounds on starting times

- Disjunctive graph given current order π
- Nodes are vehicle indices $\mathcal{N}$
- Edges $i \xrightarrow{w(i,j)} j$
  - Conjunctive edges $i \xrightarrow{\rho} j$
  - Disjunctive edges $i \xrightarrow{\sigma} j$ or $j \xrightarrow{\sigma} i$
- Lower bounds LB_π on starting times given current order π

$$\mathrm{LB}_\pi(j) = \max\{r_j, \mathrm{LB}_\pi(i) + w(i,j)\}$$

Imitation learning with neural policy



- crossing order $\pi = ((1,1),(2,1))$ of vehicles
- step-by-step construction of this order
  - 1. choose $(1,1)$
  - 2. choose $(2,1)$
  - 3. $\ldots$

Imitation learning with neural policy

- get optimal trajectories from MILP solver
- parameterize policy based on LB_π
  - only consider $\mathrm{LB}_\pi(j)$ for unscheduled j
  - recurrent embedding of $\mathrm{LB}_\pi(j)$ per lane
  - alternatively, use zero padding
- fit policy parameters to expert transitions

Lower-level trajectory optimization

- position x, velocity v, control input u
- position of vehicle in front x', follow distance L
- position of intersection B, crossing time τ

$$
\begin{aligned}
\arg\min_{x:[0,\tau]\to\mathbb{R}} \int_0^\tau |x(t)|dt & & \\
\text{s.t. } \ddot{x}(t) = u(t), & & \text{for all } t \in [0,\tau], \\
|u(t)| \le a_{\max}, & & \text{for all } t \in [0,\tau], \\
0 \le \dot{x}(t) \le v_{\max}, & & \text{for all } t \in [0,\tau], \\
x'(t) - x(t) \ge L, & & \text{for all } t \in [0,\tau], \\
(x(0), \dot{x}(0)) = s_0, & & \\
(x(\tau), \dot{x}(\tau)) = (B, v_{\max}) & &
\end{aligned}
$$