8. High-Speed Signal-Free Intersection II

金力 Li Jin
li.jin@sjtu.edu.cn
上海交通大学密西根学院
Shanghai Jiao Tong University UM Joint Institute

上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

Recap

- Macroscopic decisions for platooning
- Link level: headway regulation
  - Hybrid fluid model
  - Analysis
  - Design
- Junction level: cooperative scheduling
  - Static approach
  - Dynamic approach
- Network level: cooperative routing
  - Problem formulation
  - Fundamental tradeoff

Outline

- Sequencing problem
  - Hierarchical control system
  - Model formulation
- Stability analysis
  - Stochastic stability
  - Stability of Markov processes*
  - Theoretical results*
- Optimality analysis
  - Course project guidelines

Sequencing problem

- Hierarchical control system
- Model formulation

Hierarchical control system

- Upper level: sequencing
  - A centralized controller (e.g. RSU) determines the sequence of CAVs
  - Sequencing leads to time windows for each CAV to cross
- Lower level: trajectory planning
  - A CAV plans its trajectory to ensure crossing during the allocated time window (absorbing delay en route)
  - Vehicle following or coordination needed

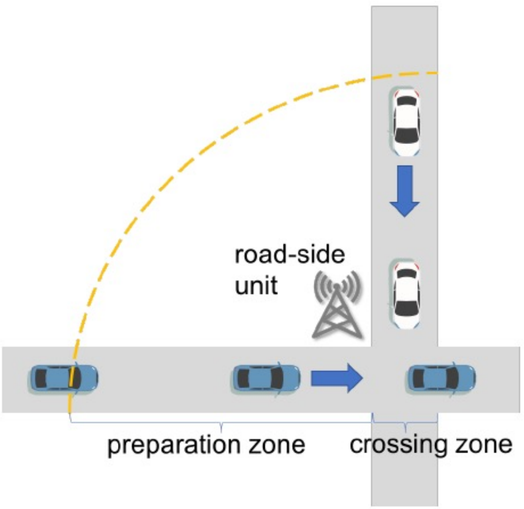
sequencing

↓

trajectory planning

Sequencing problem

- Consider two CAVs (labeled 1 & 2) consecutively crossing the intersection
- Suppose CAV k crosses the intersection at time t_k
- What constraints are imposed on the crossing times?
  - If vehicle k enters the control zone at time s_k, it cannot cross until $s_k + \Delta$, where Δ is the nominal traverse time.
  - The crossing times t_1, t_2 should be staggered; no less than θ



sequencing

↓

trajectory planning

Sequencing problem

- How to determine safe headway θ?
- Simple case:
  - CAVs 1 & 2 are from the same direction
  - Minimal headway: as in platooning
- More complex case #
  - CAVs 1 & 2 on intersecting orbits: safety constraint
  - CAVs 1 & 2 on non-intersecting orbits: no constraint

Problem formulation

- We start with a simplified, discrete-time formulation.
- Two one-way orthogonal orbits without turning
- State: $X_k(t) =$ # of CAVs waiting in direction k
- At each time step, a CAV arrives in direction k with probability $p_k \in [0,1]$
- At each time step, the intersection can discharge at most one CAV
  - Same-direction headway $\theta_{11} = \theta_{22} = 1$ [time step]
  - Orthogonal-direction headway $\theta_{12} = \theta_{21} = 2$ [time steps] (very fake number)
  - We can use an auxiliary dummy variable to formulate it

System dynamics

- A bit complex; fasten you seatbelt…
- System state $X(t) = [X_1(t), X_2(t)]^T \in \mathbb{Z}^2_{\geq 0}$
  - $X_k(t)$ = # of CAVs waiting in direction k
- The above is not enough, we need an auxiliary state $Y = \{0,1,2\}$, which is the "previous vehicle class"
  - $Y(t) = k$ if a class-k vehicle was discharged at $t-1$
  - $Y(t) = 0$ if no vehicle was discharged at $t-1$
- Action $A(t) \in \{1,2\}$
  - $A(t) = k$ essentially (but not exactly) means direction k is being discharged at time t

Discharging mechanism

- Let Δ_k be the # of vehicles arriving in direction k
- $\Delta_k = \begin{cases} 1 & \text{with probability } p_k \\ 0 & \text{with probability } 1 - p_k \end{cases}$
- $X_k(t+1) =$
$\begin{cases} (X_k(t) + \Delta_k - 1)_+ & \text{if } A(t) = k \text{ and } if\ Y(t) = k \text{ or } 0 \\ X(t) + \Delta_k & \text{otherwise} \end{cases}$
- $(\cdot)_+$ represents the positive part of a function

$$(\xi)_+ = \begin{cases} \xi & \xi \geq 0 \\ 0 & o.w. \end{cases}$$

- How are queues discharged?

Discharging mechanism

- If we discharged a class-1 vehicle at time t, we can immediately discharge another class-1 vehicle at time $t+1$
- If we discharged a class-1 vehicle at time t, we cannot discharge a class-2 vehicle until time $t+2$
- If no vehicle is discharged at time t, we can discharge a vehicle of either class at time $t+1$
  - Case 1: empty intersection
  - Case 2: switching over

Transition probabilities

- $p(x', y'|x, y, a)$ = probability that $X(t+1) = x'$, $Y(t+1) = y'$ conditional on $X(t) = x, Y(t) = y$, $A(t) = a$.
- Notational convention:
  - Capital letter = random variables: X, Y, A
  - Lower-case letter = numbers x, y, a
  - CDF $F_X(x)$, PMF $p_X(x)$, PDF $f_X(x)$
  - Never write "$\Pr\{x = 1\}$" or "$f(X)$ is increasing in X"
- Suppose that $X(t) = [2,3]^T$, $Y(t) = 1$, and $A(t) = 1$.
  - Pr{$X(t+1) =*, Y(t+1) =*$}=?
  - Pr{[1,3],1}=(1-p1)(1-p2), Pr{[2,3],1}=p1(1-p2), Pr{[1,4],1}=(1-p1)p2, Pr{[2,4],1}=p1p2

Transition probabilities

- Suppose that $X(t) = [2,3]^T$, $Y(t) = 1$, and $A(t) = 2$.
  - Pr{[2,3],0}=(1-p1)(1-p2), Pr{[3,3],0}=p1(1-p2), Pr{[2,4],0}=(1-p1)p2, Pr{[3,4],0}=p1p2
- Suppose that $X(t) = [2,3]^T$, $Y(t) = 0$, and $A(t) = 1$.
  - Pr{[1,3],1}=(1-p1)(1-p2), Pr{[2,3],1}=p1(1-p2), Pr{[1,4],1}=(1-p1)p2, Pr{[2,4],1}=p1p2
- Suppose that $X(t) = [0,3]^T$, $Y(t) = 1$, and $A(t) = 1$.
  - Pr{[0,3],0}=(1-p1)(1-p2), Pr{[0,3],1}=p1(1-p2), Pr{[0,4],0}=(1-p1)p2, Pr{[0,4],1}=p1p2
- Suppose that $X(t) = [0,0]^T$, $Y(t) = 1$, and $A(t) = 1$.
  - Pr{[0,0],0}=(1-p1)(1-p2), Pr{[0,0],1}=p1(1-p2), Pr{[0,1],0}=(1-p1)p2, Pr{[0,1],1}=p1p2
- Suppose that $X(t) = [0,0]^T$, $Y(t) = 2$, and $A(t) = 1$.
  - Pr{[0,0],0}=(1-p1)(1-p2), Pr{[1,0],0}=p1(1-p2), Pr{[0,1],0}=(1-p1)p2, Pr{[1,1],0}=p1p2

A quick note on dynamic programming

The process $\{X(t), Y(t); t = 0,1,2, \dots\}$ is a discrete-time, discrete-state Markov process.

- Markov processes
- Markov decision processes
- Agent-environment interface
- Reference: Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT Press, 2018. http://incompleteideas.net/book/RLbook2020.pdf (Optional)

Markov process

- Stochastic process: random variables evolving over time
  - Flip a coin
  - Count vehicles
  - Measure power demand
- Mathematically, we use a time-varying random variable S_t to describe a stochastic process
- $\Pr\{S_t = s | S_{t-1}, \dots S_1, S_0\}$
- Markov process: the distribution of S_t only depends on S_{t-1} and does not depend on $S_0, \dots, S_{t-2}$
- $\Pr\{S_t = s | S_{t-1}, \dots S_1, S_0\} = \Pr\{S_t | S_{t-1}\}$

Андре́й А. Ма́рков
Andrey A. Markov
安德烈·A·马尔可夫
1856-1922

Markov decision process

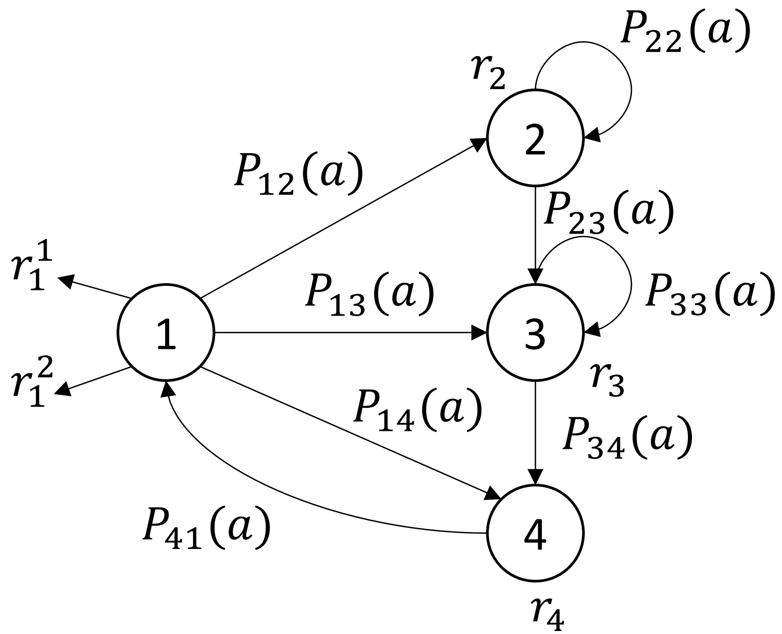
- Markov decision process: at each time, we can take some action that affects the evolution of the stochastic process
- Example:
  - Times at which vehicles enter and pass an intersection is random
  - But we can influence these random variables by controlling the traffic signal
  - Mathematically, the traffic signal control action will affect the distribution of the random enter/pass times

Agent-environment interface

- Agent-environment interaction in a MDP

- MDP trajectory:
  - Time sequence t = 0, 1, 2, …
  - State: $S_t \in \mathcal{S}$
  - Action: $A_t \in \mathcal{A}(s)$
  - Reward: $R_t \in \mathcal{R}$
  - Trajectory: $S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, \ldots$

Graphic representation



Question: what if the action a only depends on the current state s?

-> an ordinary Markov chain (as we have seen in the previous lecture)!

Dynamics

- We use function p to describe dynamics of MPD:
$p(s',r|s,a) := \Pr\{S_t = s', R_t = r | S_{t-1} = s, A_{t-1} = a\}$
- Reward may or may not be random

- With a slight abuse of notation, state-transition probabilities

$$p(s'|s,a) := \Pr\{S_t = s' | S_{t-1} = s, A_{t-1} = a\}$$
$$= \sum_{r \in \mathcal{R}} p(s',r|s,a)$$

Reward

- Expected reward

$$r(s,a) \coloneqq \mathrm{E}\{R_t | S_{t-1} = s, A_{t-1} = a\}$$
$$= \sum_{r\in\mathcal{R}} r \sum_{s\prime\in\mathcal{S}} p(s', r | s, a)$$

- Three-argument function

$$r(s,a,s') \coloneqq \mathrm{E}\{R_t = r | S_{t-1} = s, A_{t-1} = a, S_t = s'\}$$
$$= \sum_{r\in\mathcal{R}} r \frac{p(s', r | s, a)}{p(s' | s, a)}$$

MDP formulation

- State $S(t) = \big(X(t), Y(t)\big) \in \mathbb{Z}_{\geq 0}^2 \times \{0,1,2\}$
- Action $A(t) \in \{1,2\}$
- Dynamics (transition probabilities)

$$p(s'|s, a) = p(x', y'|x, y, a)$$

- Reward $R(t) = -\|X(t)\|_1 = -X_1(t) - X_2(t)$
- Return $G(t) = \sum_{s=t}^{\infty} \gamma^{s-t} R(s)$

Sequencing policy

- First come first serve (FCFS)
  - Fair, easy
- Minimal switch-over (MSO)
  - Efficient, but maybe unfair
- Longer queue first (LQF)
  - Fairer
- Two metrics for evaluation
  - Throughput: maximal demand that the intersection can accommodate
  - Waiting time: Queuing delay experienced by vehicles

Stability analysis

- Stochastic stability
- Stability of Markov processes*
- Theoretical results*

Stochastic stability

- The intersection is stochastically stable if there exists $Z < \infty$ such that for any initial condition $(x, y) \in \mathbb{Z}_{\geq 0}^2 \times \{0,1,2\}$, we have

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t} \mathrm{E}[X_1(t) + X_2(t)] \leq Z.$$

- Interpretation: time-average is bounded
- Stronger stability:

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t} \mathrm{E}\left[X_1^p(t) + X_2^p(t)\right] \leq Z$$

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t} \mathrm{E}\left[\exp\left[\rho\left(X_1(t) + X_2(t)\right)\right]\right] \leq Z$$

Stability of Markov processes*

- For Markov processes, stability involves two notions.
- Boundedness: some moment of the state $S(t)$ is bounded on average.

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t} \mathrm{E}[|S(t)|^p] \le Z$$

- Convergence: distribution of state $P_t(s)$, conditional on $S(0) = s$, converges to a unique steady-state distribution P.

$$\lim_{t\to\infty} \|P_t(s) - P\|_{\mathrm{TV}} = 0$$

(TV = total variation distance)

Theoretical result*

- Very advanced theory (PhD level); fasten your seatbelt!
- Let's start with the MSO policy (for ease of presentation)
- Recall that our objective is to show
$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t} \mathrm{E}[X_1(t) + X_2(t)] \leq Z.$$
- Direct evaluation of the limit in the left-hand side (LHS) is not easy
- Instead, we consider a Lyapunov function
$$V(x) = \frac{1}{2}(x_1 + x_2)^2$$

Алекса́ндр М. Ляпуно́в
Aleksandr M. Lyapunov
亚历山大·M·李亚普诺夫
1857-1918

Theoretical result*

- We use a tool called Foster-Lyapunov theorem: A Markov process $S(t)$ is stable if there exists $c > 0$ and $d < \infty$ such that
$$\mathrm{E}\left[V\left(S(t+1)\right) - V\left(S(t)\right) \middle| S(t) = s\right] \leq -cs + d$$
for all s in the state space.
- The above ensures that
$$\limsup_{t\to\infty} \frac{1}{t}\sum_{s=0}^{t} \mathrm{E}[S(t)] \leq \frac{d}{c}$$
- Now, the question is how to evaluate the expected increment $\mathrm{E}\left[V\left(S(t+1)\right) - V\left(S(t)\right) \middle| S(t) = s\right]$

Theoretical result*

- Law of total expectation: let $B_1, B_2, \ldots, B_n$ be a partition of the sample space, i.e., a set of disjoint events that collectively cover the whole sample space. Then,

$$\mathrm{E}[A] = \sum_{i=1}^{n} \mathrm{E}[A|B_i] \Pr\{B_i\}$$

- Recall that for the intersection model, we have the transition probabilities $p(x', y'|x, y, a)$
- So, we have

$$E\left[V\big(X(t+1)\big) - V\big(X(t)\big)|X(t) = x, Y(t) = y\right]$$
$$= \sum_{x' \in \mathbb{Z}_{\geq 0}^2} p(x', y'|x, y, a)\, V(x') - V(x)$$

Theoretical result*

- One can plug the expression for $p(x', y'|x, y'a)$ into the above equation and construct c, d verifying the Foster-Lyapunov theorem.
- After very involved math, we can conclude that the MSO policy stabilizes the intersection if and only if

$$p_1 + p_2 < 1$$

- That is, MSO is stabilizing if and only if demand <supply
- Hence, no policy can accommodate a higher demand (or attain a higher throughput) than MSO
- Hence, MSO maximizes throughput.

Optimality analysis

- Course project guidelines

Possible topic for course project

- Find the optimal sequencing policy
- Reward = $-\left(X_1(t) + X_2(t)\right) - |X_1(t) - X_2(t)|$
  - Balance between efficiency and fairness
- Use either Monte-Carlo or temporal-difference method
- Reference: 2. Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT Press, 2018. http://incompleteideas.net/book/RLbook2020.pdf

Summary

- Sequencing problem
  - Hierarchical control system
  - Model formulation
- Stability analysis
  - Stochastic stability
  - Stability of Markov processes*
  - Theoretical results*
- Optimality analysis
  - Course project guidelines

Next time

- 6/4: Guest lecture by Prof. Junyu Cao from UT Austin, optimization of smart delivery service
  - Assistant Professor, UT Austin
  - PhD, UC Berkeley
  - BE, Xi’an Jiao Tong University (XJTU)
- 6/7: Adaptive ramp metering
  - HW2 due
  - Preliminary project teaming due
- 6/16: Quiz 1
  - 60 min in class (8:15-9:15AM)
  - Online students must turn on video
  - Open lecture notes, but nothing else
  - Will have a review session