AA 598B Special Topics

Decision-Making & Control for Safe Interactive Autonomy

Instructor: Prof. Karen Leung

Autumn 2024

https://faculty.washington.edu/kymleung/aa598/

Announcements

- Guest lecture on Wednesday by Dr. Boris Ivanovic, Senior Research Scientist and Manager in NVIDIA Autonomous Vehicle Research Group
  - Submit talk review/reflection
- Homework 2 out tomorrow (will be light)
- Start thinking about project proposals
  - Due Nov 1 Friday
- Next Wednesday long paper discussion
- Useful video from IROS: https://www.youtube.com/watch?v=QYbAvOPcy0s

Project proposal

- **Research project**—A research project that is connected to topics covered in this course. It is encouraged for it be to connected to your PhD/MS research or other course projects. But the connection to the course must be evident and the contributions distinct.
- **Literature survey**—A deep dive into several papers on a chosen topic area, including your inclusion criteria, motivating questions, and insights.
- The grading for the project is as follows.
  - Project proposal due week 5 (5%)
  - Project presentation in week 10 (10%)
  - Project presentation peer review in week 10 (5%)
  - Project report due finals week (15%)

Last time

- Wrapped up behavior prediction models
  - Ontological “theory of mind” approaches
  - Phenomenological “deep generative model” approaches
  - Many different datasets and open-source code available
  - Many different metrics are used to evaluate prediction performance
  - Prediction models are being adapted for generate realistic human behaviors for simulators
    - Still a hard problem (stability, controllability, multi-agents, etc)

Today

- Start planning module!
  - Defining the problem
  - Techniques to solve the problem

Interaction-aware planning

Module #2


Traffic in the right lane is exiting a freeway.

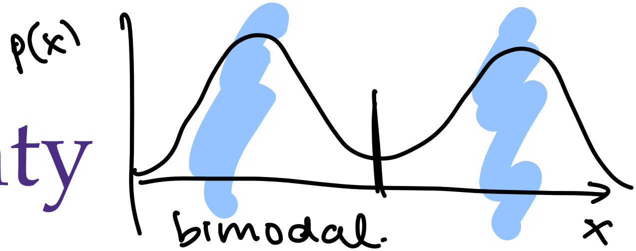
Goal: Compute $u_{robot} = \pi(x_{robot}, x_{humans}, e)$

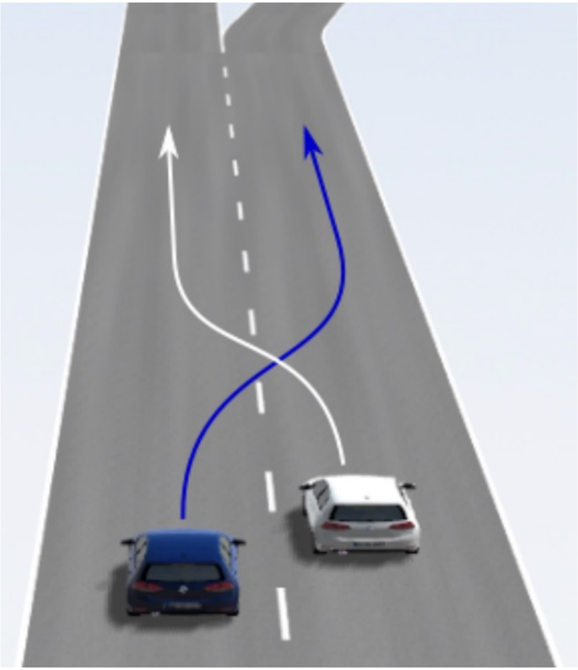
What makes interaction-aware planning challenging?

(discussion)

- multiple agent interaction → combinatorial explosion → hard to find tractable planning alg.
  - selecting a view cone.
- sensing/tracking multiple agents can be hard. → affect planning frequency
- uncertainty in human tendencies. → models for uncertainty, ideally avoid strong assumptions.
- conservative vs efficiency
- accounting for out of distribution events.
- adapting to new/unseen settings./OOD out of dis.
- planning horizon?
- feedback from human ⇔ robot
  - behaviors may amplify

Multimodal uncertainty

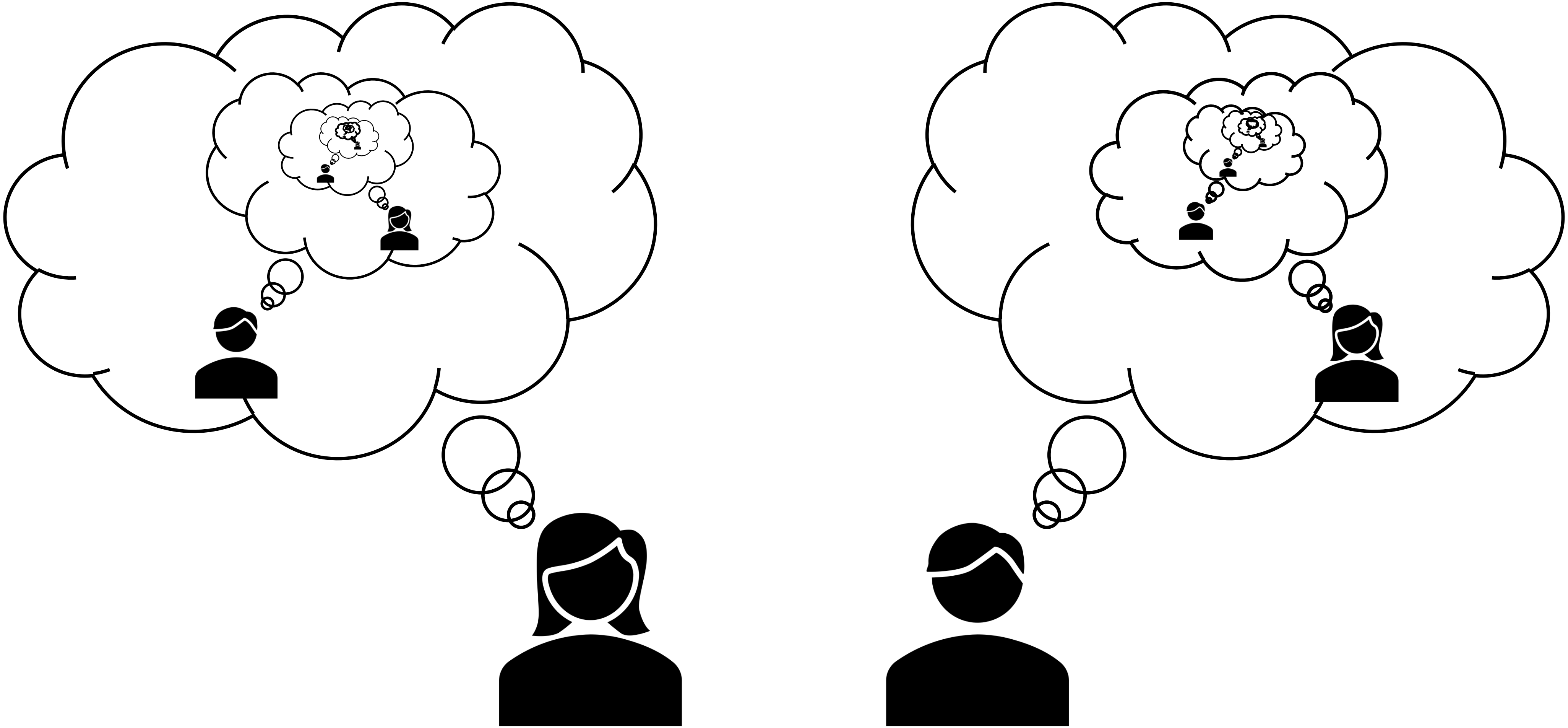




Schmerling et al 2018

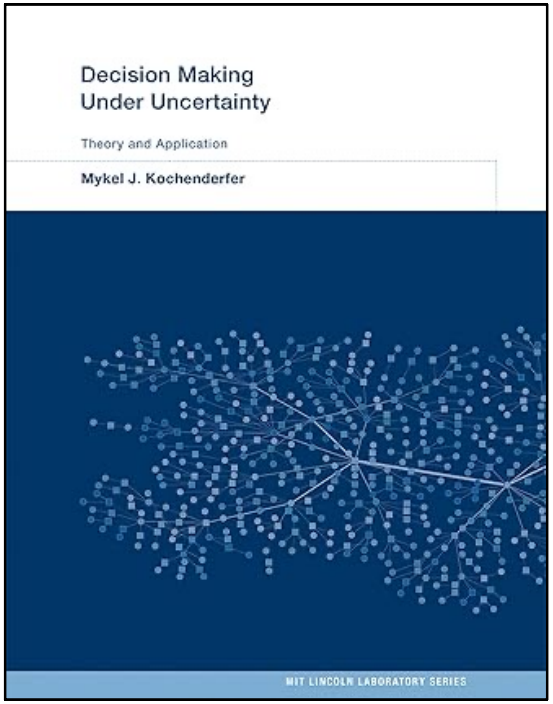
GMM: gaussian mixture model

- the mean is not a good measure of outcome.
- outcomes can be very distinct → lead to very different plans.

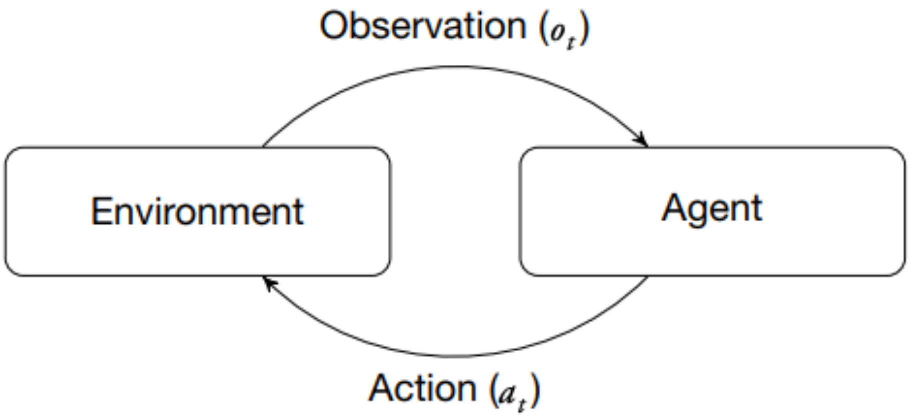


Need to account for how others may respond to your own actions

Decision-making under uncertainty



https://mykel.kochenderfer.com/textbooks/



Planning problem: Find **actions** that accomplish the desired **task**

- Actions: How are actions represented?

- next robot state for the next timestep
  + tracking controller (PID, LQR)
- control inputs. $u_R = \pi(x_R, \ldots)$
- desired trajectory / sequence of waypoints + tracking controller.

Planning problem: Find **actions** that accomplish the desired **task**

Task: How is the task defined?

- objective functions + constraints → mathematical functions describing these.

- indicator function for task success / failure.
- demonstrations of success / failure
- LLMs / define through language.
- temporal logic, a formal language to express specifications.
  STL, LTL

Approaches to solving a planning problem

- **Search-based**: Enumerate over all possible options and pick the best one BFS, DFS, A*, sampling-based motion planning
  PRM, RRT*, FMT*

Approaches to solving a planning problem

- **Search-based**: Enumerate over all possible options and pick the best one
- **Supervised learning**: Mimic what an expert did (i.e., behavior cloning)

Approaches to solving a planning problem

- **Search-based**: Enumerate over all possible options and pick the best one
- **Supervised learning**: Mimic what an expert did (i.e., behavior cloning)
- **Optimization-based**: Assume problem dynamics and frame as optimization problem fmincon, cvxpy, IPOPT...

Approaches to solving a planning problem

- **Search-based**: Enumerate over all possible options and pick the best one
- **Supervised learning**: Mimic what an expert did (i.e., behavior cloning)
- **Optimization-based**: Assume problem dynamics and frame as optimization problem
- **Reinforcement learning**: Learn from interactions & optimize

CS599 ~~G~~ Abhishek Gupta.

General optimal planning problem

(assume no human) $x_R^{(t)} \in \mathbb{R}^n$, $u_R^{(t)} \in \mathbb{R}^m$, $t \in \mathbb{R}_+$

goal is to find $u_R^{(0)}, \ldots u_R^{(T)}$ that accomplishes the task.

$$\min_{u_R^{(0:T)}} J\left(x_R^{(0:T+1)}, u_R^{(0:T)}\right)$$

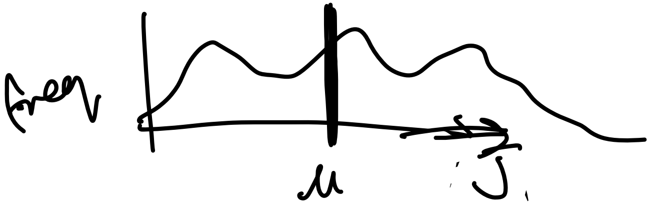
s.t. $x_R^{(t+1)} = f\left(x_R^{(t)}, u_R^{(t)}\right)$ dynamics

$g_i\left(x_R^{(0:T+1)}, u_R^{(0:T)}\right) \geq 0 \quad i = 1 \ldots, G$ inequality eg. control limits obstacles, speed ..

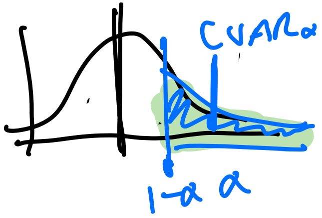
$h_j\left(x_R^{(0:T+1)}, u_R^{(0:T)}\right) = 0 \quad j = 1 \ldots, H$ equality initial state. goal state

eg. STL.

with human(s) CVAR



$$\min_{u_R^{(0:T)}} \mathbb{E}_{u_H \sim p(u_H \dots)} \left[J\left(x_R^{(0:T+1)}, u_R^{(0:T)}, x_H^{(0:T+1)}, u_H^{(0:T)}\right) \right]$$

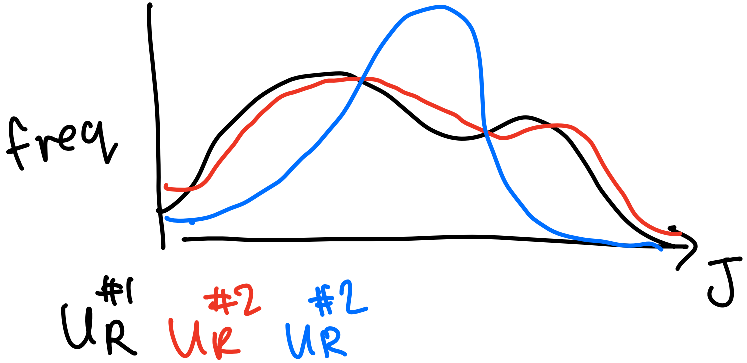


s.t. $x_R^{(t+1)} = f_R\left(x_R^{(t)}, u_R^{(t)}\right), \quad x_H^{(t+1)} = f_H\left(x_H^{(t)}, u_H^{t}\right)$

$$g_i\left(x_R^{(0:T+1)}, u_R^{(0:T)}, x_H^{(0:T+1)}, u_H^{(0:T)}\right) \geq 0 \quad i = 1, \dots, G$$

$$h_j\left(x_R^{(0:T+1)}, u_R^{(0:T)}, x_H^{(0:T+1)}, u_H^{(0:T)}\right) = 0 \quad j = 1, \dots, H.$$

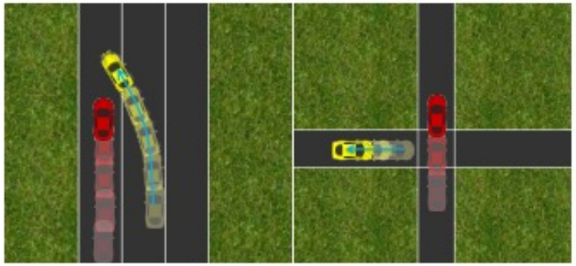
but we have $u_H^{(t)} \sim p\left(u_H \mid x_R^{(t)}, x_H^{(t)}, u_R^{(t)}, \dots\right)$ or $x_H^{(t)}$ $\quad u_H^{(0:T)}$ is very messy.



$$D\left(p(J \mid u_R^{\#1}) \,\|\, p(J \mid u_R^{\#2})\right)$$

Stochastic optimal control

Planning with access to other agent's reward



(a) Car merges *ahead* of human; anticipates human *braking*

(b) Car *backs up* at 4way stop; anticipates human *proceeding*

(c) User drives human car

Fig. 1: We enable cars to plan with a model of how human drivers would react to the car's actions. We test the planner in a user study, where the car figures out that (a) it can merge in front of a human and that will slow them down, or (b) it can back up slightly at an intersection and that will make the human go first.

At every iteration, the robot needs to find the $\mathbf{u}_\mathcal{R}$ that maximizes this reward:

$$\mathbf{u}_\mathcal{R}^* = \arg\max_{\mathbf{u}_\mathcal{R}} R_\mathcal{R}(x^0, \mathbf{u}_\mathcal{R}, \mathbf{u}_\mathcal{H}^*(x^0, \mathbf{u}_\mathcal{R})) \qquad (5)$$

Here, $\mathbf{u}_\mathcal{H}^*(x^0, \mathbf{u}_\mathcal{R})$ is what the human would do over the next N steps if the robot were to execute $\mathbf{u}_\mathcal{R}$.

The robot does not actually know $\mathbf{u}_\mathcal{H}^*$, but in the next section we propose a *model* for the human behavior that the robot can use, along with an approximation to make (5) tractable.

$$\mathbf{u}_\mathcal{H}^*(x^0, \mathbf{u}_\mathcal{R}) = \arg\max_{\mathbf{u}_\mathcal{H}} R_\mathcal{H}(x^0, \mathbf{u}_\mathcal{R}, \mathbf{u}_\mathcal{H})$$

Sadigh et al 2016

Planning with access to other agent's reward

To apply L-BFGS, we need the gradient of (5) with respect to $\mathbf{u}_{\mathcal{R}}$:

$$\frac{\partial R_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{R}}} = \frac{\partial R_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{H}}} \frac{\partial \mathbf{u}^*_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{R}}} + \frac{\partial R_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{R}}} \tag{11}$$

$\frac{\partial R_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{H}}}$ and $\frac{\partial R_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{R}}}$ can both be computed symbolically through backward propogation, as we have a representation of $R_{\mathcal{R}}$ in terms of $\mathbf{u}_{\mathcal{H}}$ and $\mathbf{u}_{\mathcal{R}}$. For $\frac{\partial \mathbf{u}^*_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{R}}}$, we use that $\mathbf{u}^*_{\mathcal{H}}$ is the minimum from (10), which means that the gradient of $R_{\mathcal{H}}$ evaluated at $\mathbf{u}^*_{\mathcal{H}}$ is 0:

$$\frac{\partial R_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{H}}}(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}^*_{\mathcal{H}}(x^0, \mathbf{u}_{\mathcal{R}})) = 0 \tag{12}$$

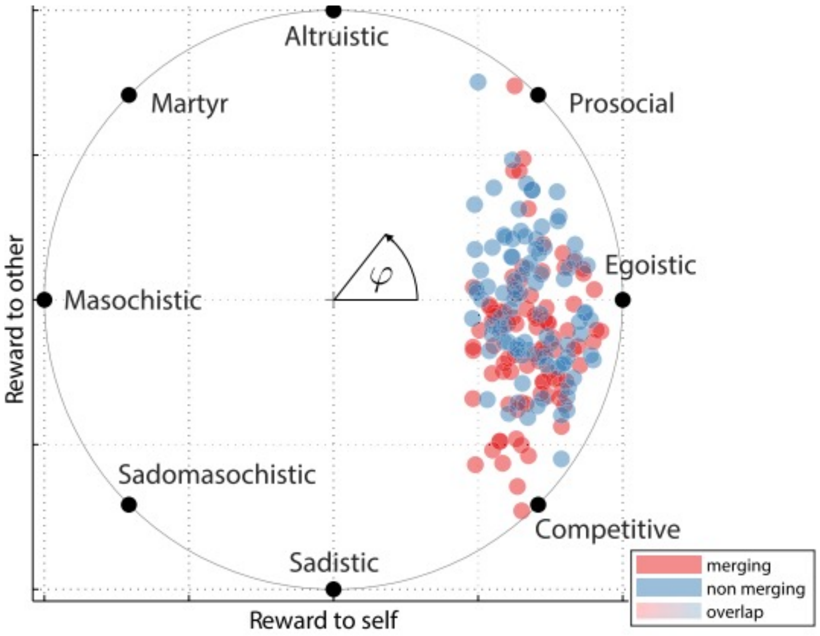
Now, we can differentiate the expression in equation (12) with respect to $\mathbf{u}_{\mathcal{R}}$:

$$\frac{\partial^2 R_{\mathcal{H}}}{\partial \mathbf{u}^2_{\mathcal{H}}} \frac{\partial \mathbf{u}^*_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{R}}} + \frac{\partial^2 R_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{H}} \partial \mathbf{u}_{\mathcal{R}}} \frac{\partial \mathbf{u}_{\mathcal{R}}}{\partial \mathbf{u}_{\mathcal{R}}} = 0 \tag{13}$$

Finally, we can solve for a symbolic expression for $\frac{\partial \mathbf{u}^*_{\mathcal{H}}}{\partial \mathbf{u}_{\mathcal{R}}}$:

Planning with social considerations

Social behavior for autonomous vehicles

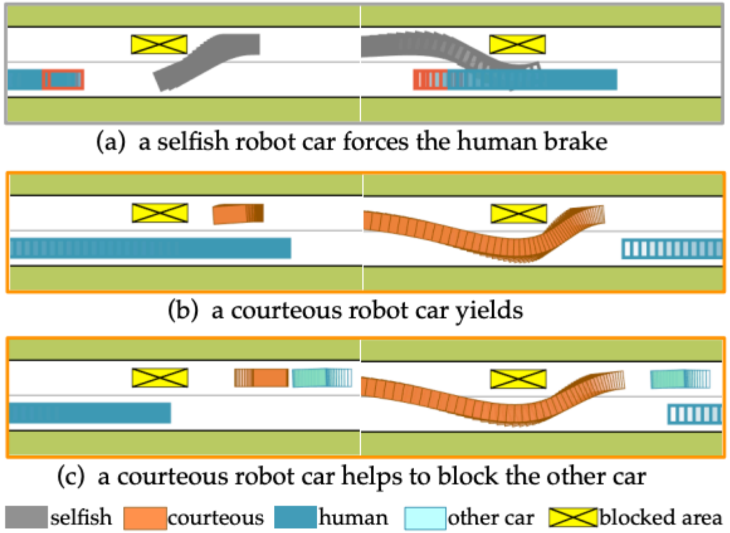


$$g_1 = \cos(\varphi_1) r_1(\cdot) + \sin(\varphi_1) r_2(\cdot),$$

Schwarting et al 2019

Planning with social considerations

Courteous Autonomous Cars



(a) a selfish robot car forces the human brake

(b) a courteous robot car yields

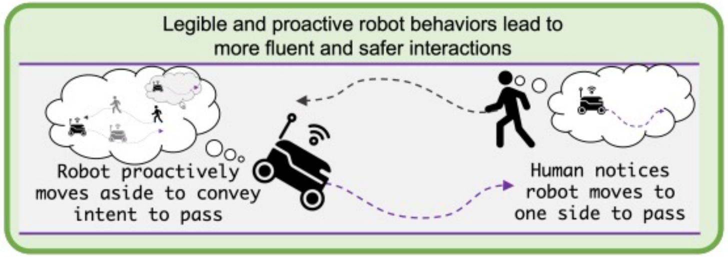
(c) a courteous robot car helps to block the other car

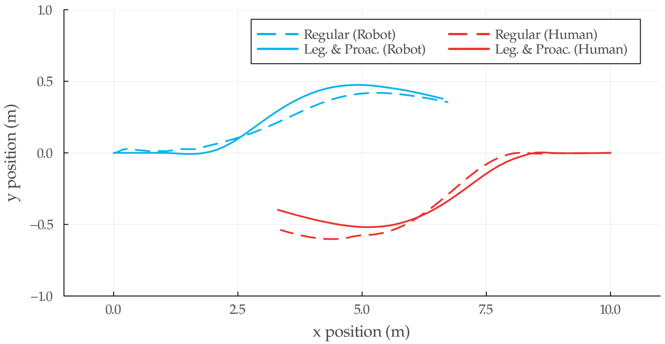
$$
\begin{aligned}
C_{\mathcal{R}}\left(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{R}}, \theta_{\mathcal{H}}, \lambda_c\right) = & \; C_{\mathcal{R}}^{self}\left(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{R}}\right) \\
& + \lambda_c \; C_{\mathcal{R}}^{court}\left(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}\right)
\end{aligned}
$$

Sun et al 2018

Planning with social considerations

Legible and Proactive Robot Planning for Prosocial Human-Robot Interactions



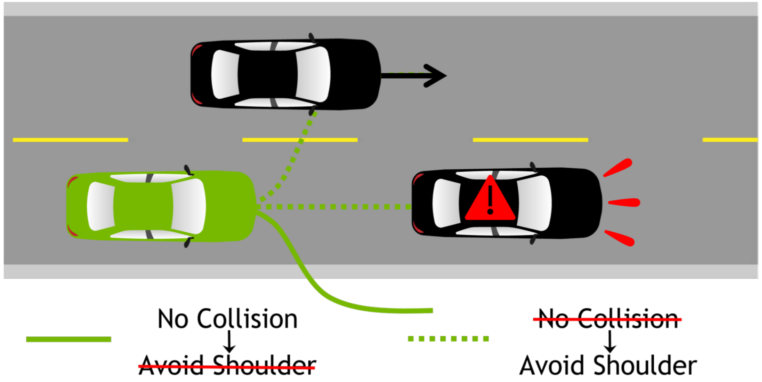


$$J_{\text{incon}}(\mathbf{x}_F^{0:T+1}, \mathbf{u}_F^{0:T}) \leq \beta_F$$

Geldenbott et al 2024

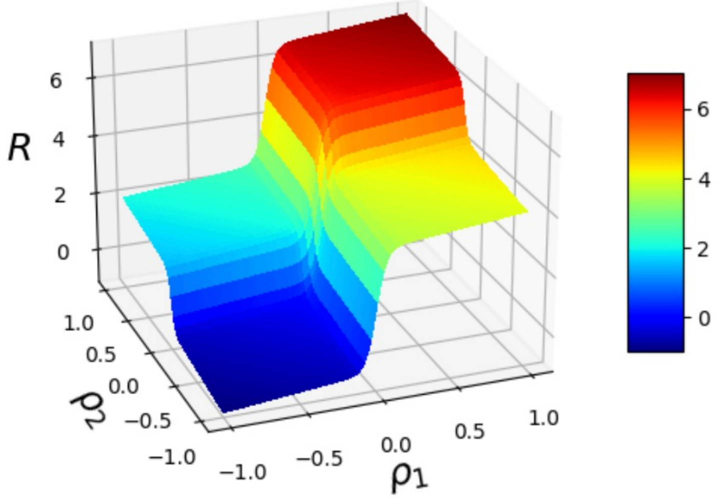
Planning with rules

Receding Horizon Planning with Rule Hierarchies for Autonomous Vehicles

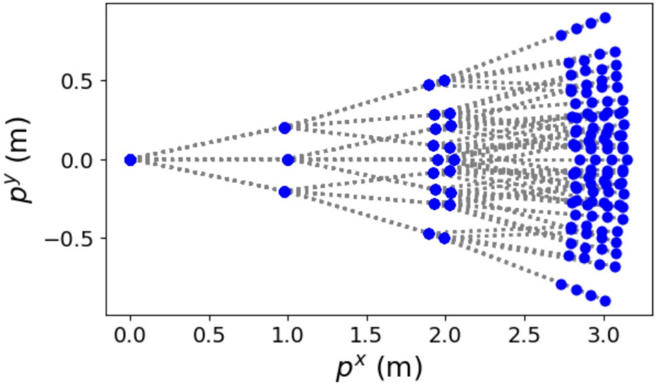


Rank	Satisfied Rules
1	ϕ_1, ϕ_2, ϕ_3
2	ϕ_1, ϕ_2
3	ϕ_1, ϕ_3
4	ϕ_1
5	ϕ_2, ϕ_3
6	ϕ_2
7	ϕ_3
8	$\emptyset$

TABLE I: Illustration of trajectory ranks for three rules.



Two-step optimization: trajectory tree + local refinement



$$R(\rho) := \sum_{i=1}^{N} \left(a^{N-i+1} \mathrm{sigmoid}(c\rho_i) + \frac{1}{N}\rho_i \right)$$

https://github.com/UW-CTRL/stljax/tree/main

What if we take advantage of parallel computation?

- So far, the methods relied on some sort of gradient descent.
- What if we can't compute gradients easily?
- We can consider searching over the space via a sampling-based approach

Planning with ego-conditioned prediction

Multimodal Probabilistic Model-Based Planning for Human-Robot Interaction

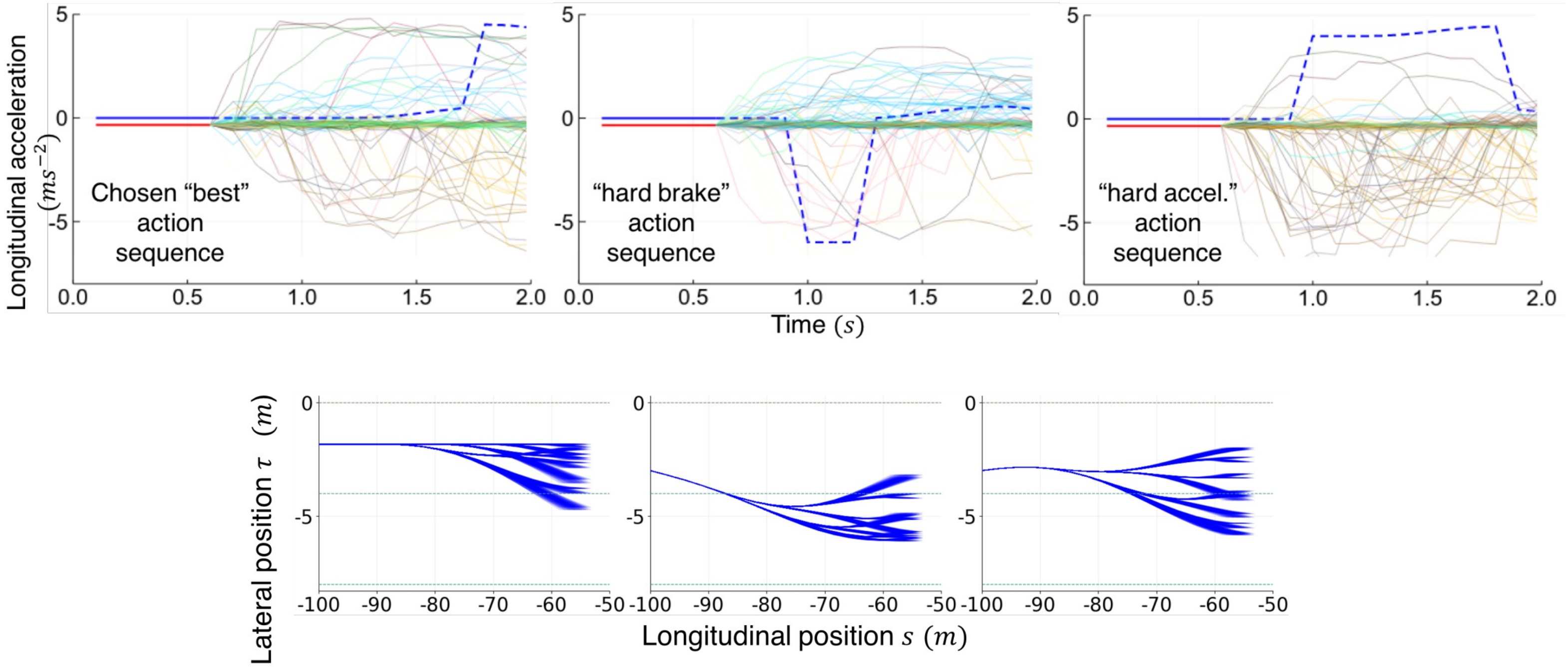


Fig. 5. 4096 candidate robot action sequences scored each planning loop.

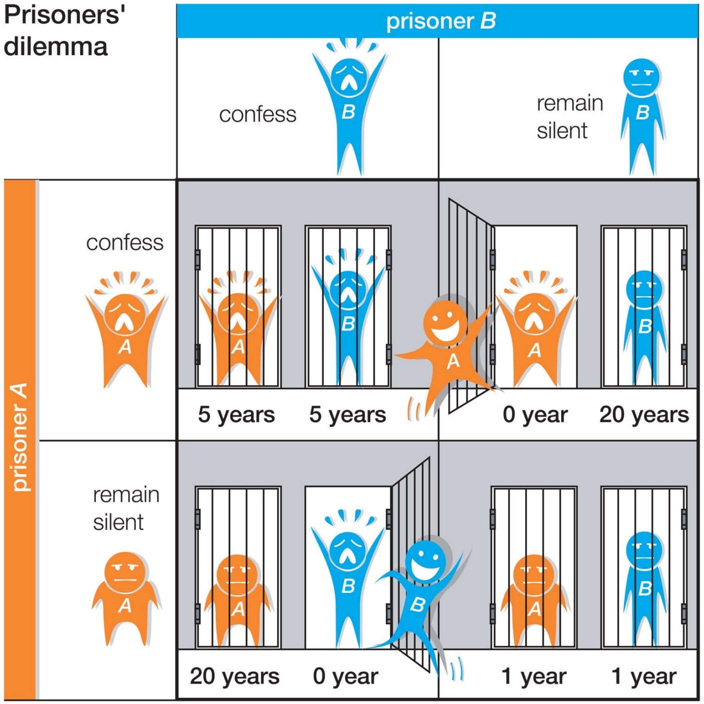
Schmerling et al 2018

Model Predictive Path Integral (MPPI)

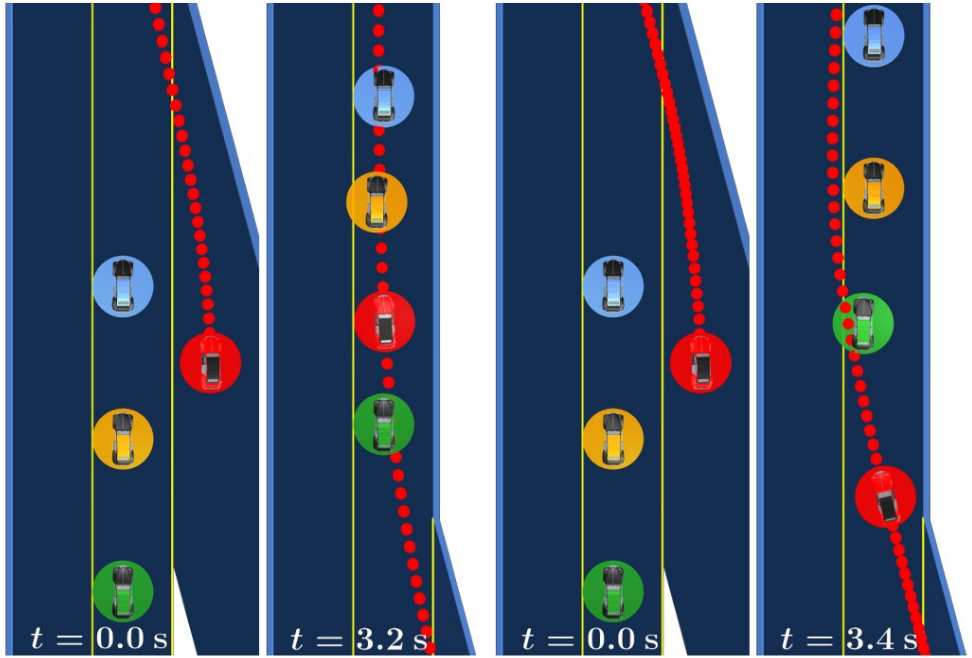
1. Start with nominal trajectory
2. Add noise to it to generate many trajectories
3. Evaluate cost of each trajectory
4. Compute weight for each trajectory
5. Compute weighted sum over controls to compute control

https://sites.gatech.edu/acds/mppi/

Game theory



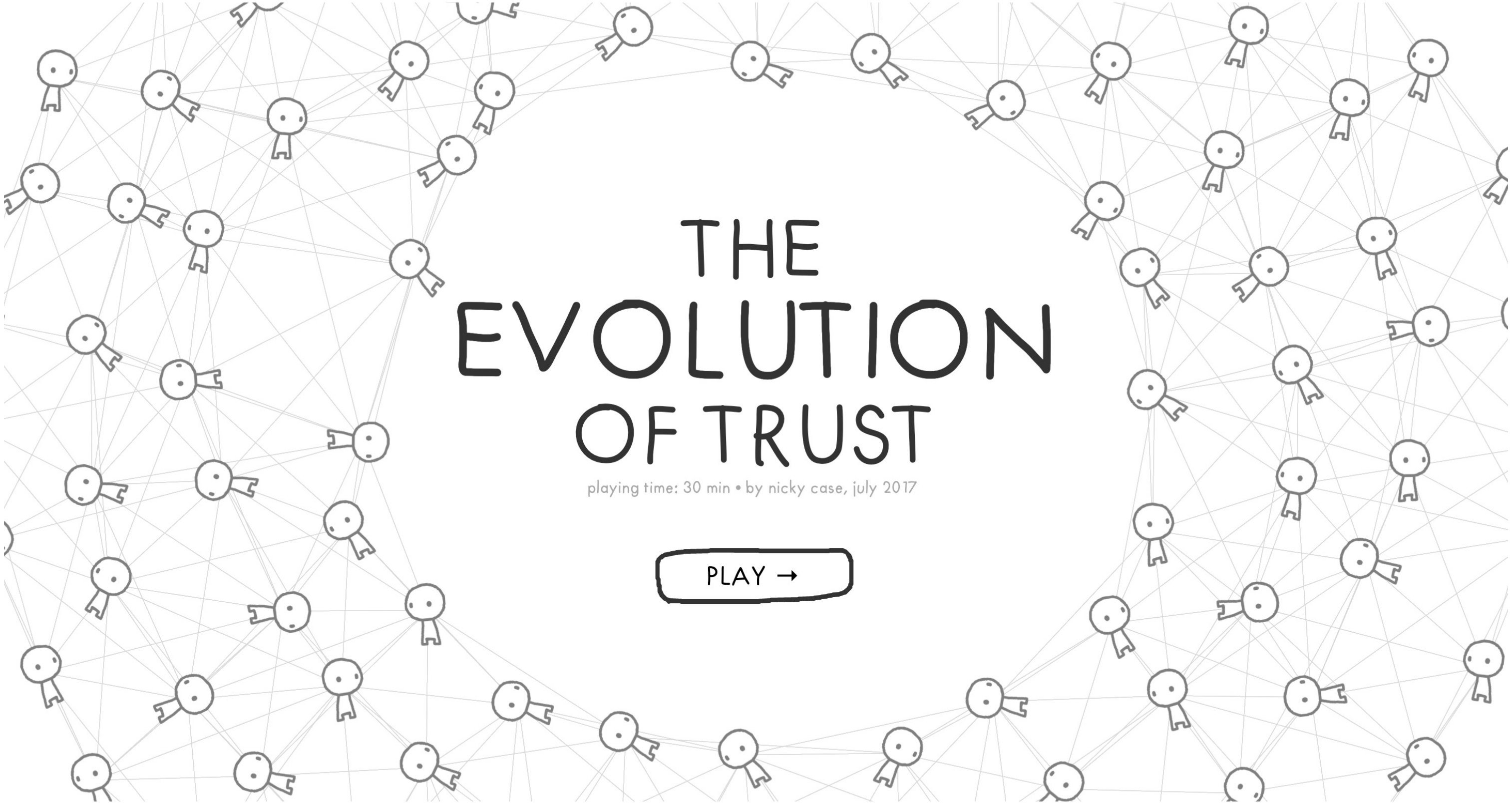
https://www.britannica.com/science/game-theory/The-von-Neumann-Morgenstern-theory



ALGAMES: A Fast Augmented Lagrangian Solver for Constrained Dynamic Games
Autonomous Robots (AuRo 2021),
S. Le Cleac'h, M. Schwager, Z. Manchester

The evolution of trust

https://ncase.me/trust/



Game theory

Definition: A mathematical framework for modeling scenarios in which multiple decision-makers (agents) interact, with each agent's outcome depending not only on its own actions but also on the actions of others.

Relevance: In human-robot interaction, game theory helps model how robots can make decisions while considering the possible actions of human agents.

General problem formulation

$$\begin{aligned} \min_{X,U^1} \quad & J^1(X,U^1) \\ \text{s.t.} \quad & D(X,U) = 0, \\ & C(X,U) \leq 0, \end{aligned} \qquad \cdots \qquad \begin{aligned} \min_{X,U^M} \quad & J^M(X,U^M) \\ \text{s.t.} \quad & D(X,U) = 0, \\ & C(X,U) \leq 0, \end{aligned}$$

$$\forall i \in [N] \begin{cases} \min\limits_{X^i, U^i} & J^i(\mathbf{X}, U^i; \theta^i) \\ \text{s.t.} & x^i_{t+1} = f^i(x^i_t, u^i_t), \forall t \in [T-1] \\ & x^i_1 = \hat{x}^i_1 \\ & {}^p g^i(X^i, U^i) \geq 0 \\ & {}^s g(\mathbf{X}, \mathbf{U}) \geq 0. \end{cases}$$

Payoff structure

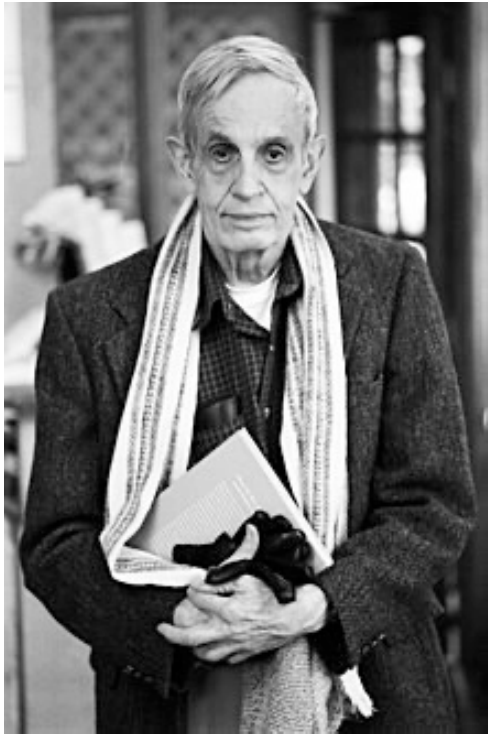
- Zero sum two player games
  - Total payoff always sums to zero
  - One player's gain is exactly equal to the other player's loss
- General sum games
  - Payoff does not need to sum to zero
  - No strong sense of win or lose

Nash equilibrium

- At Nash equilibrium, every player is playing optimally given the choices of others,
  - No player has an incentive to deviate from their chosen strategy.

$$J_i(u_i^*, u_{-i}^*) \leq J_i(u_i, u_{-1}^*) \quad \forall\, u_i \in U_i$$

- Other types:
  - Subgame Perfect Equilibrium (Nash over multiple steps)
  - Correlated Equilibrium (follow recommendation from external source)
  - Bayesian Nash Equilibrium (Nash with incomplete information, have beliefs over others)



Wikipedia