Hierarchical Game-Theoretic Planning for Autonomous Vehicles

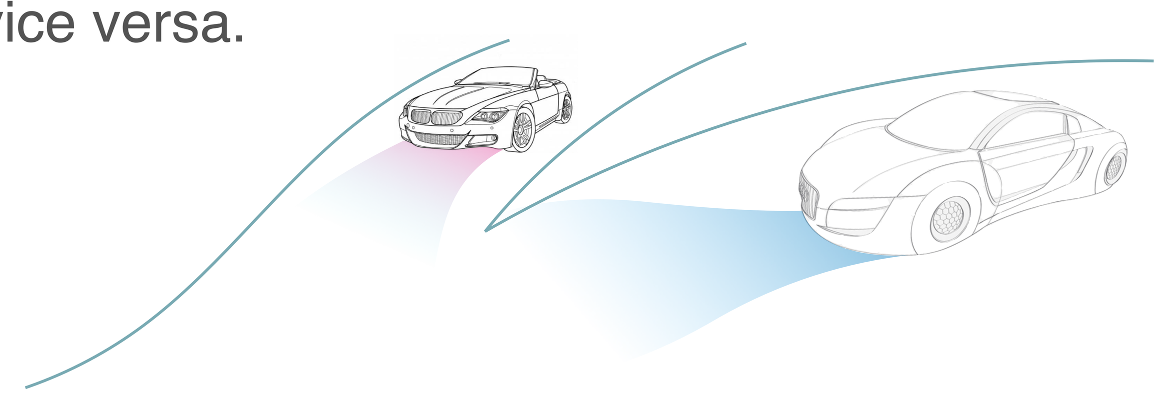
University of California, Berkeley

Jaime F. Fisac*, Eli Bronstein*, Elis Stefansson, Dorsa Sadigh, S. Shankar Sastry, and Anca D. Dragan

Road Safety: a Game-Theoretic Problem

The actions of a vehicle on the road affect and are affected by those of other drivers—autonomous vehicles are no exception.

This leads to a strong coupling of planning and prediction: others' future actions depend on our choices and vice versa.

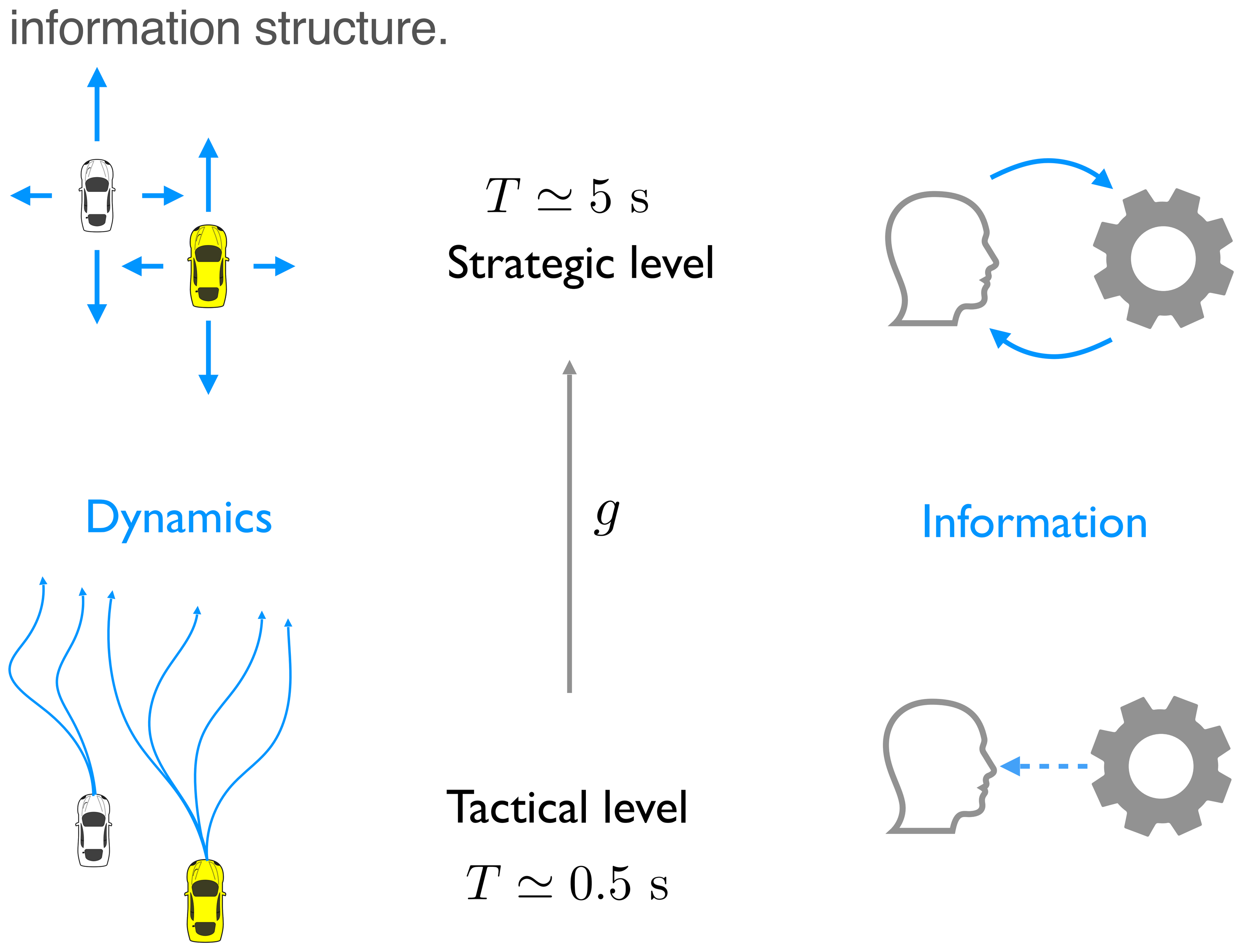


Dynamic Game Theory can reason through these interactions but is computationally challenging for high-order dynamics.

Hierarchical Game Decomposition

Our framework enables real-time tractability by structuring the dynamic game analysis into two levels.

Strategic Game: a lower-order dynamical model allows reasoning over longer horizons with a fully coupled information structure.



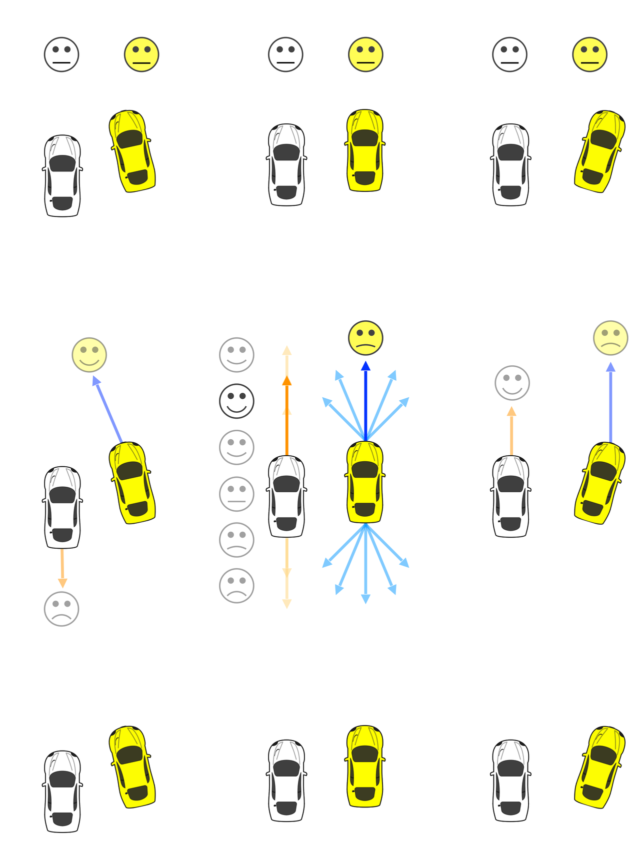
Tactical Game: the full high-order dynamical model can be used for trajectory planning with a simplified information structure.

Strategic Game: Dynamic Programming

The autonomous vehicle (A) and the human (H) reason in backward time about their ability to influence each other.

$$Q_A^k(s^k, a_A^k) = \mathbb{E}_{a_H^k} \tilde{r}_A(s^k, a_A^k, a_H^k) + Q_A^{k+1}(s^{k+1}, \pi_A^*(s^{k+1}))$$

$$Q_H^k(s^k, a_A^k, a_H^k) = \tilde{r}_H(s^k, a_A^k, a_H^k) + \mathbb{E}_{a_H^{k+1}} Q_H^{k+1}(s^{k+1}, \pi_A^*(s^{k+1}), a_H^{k+1})$$



$$\pi_A^*(s) := \arg\max_a Q_A^{k+1}(s, a) \ , \quad \forall s \in \mathcal{S}$$

$$a_H^i \sim \pi_H\left[Q_H^i(s^i, a_A^i, \cdot)\right] \ , \quad i \in \{k, k+1\}$$

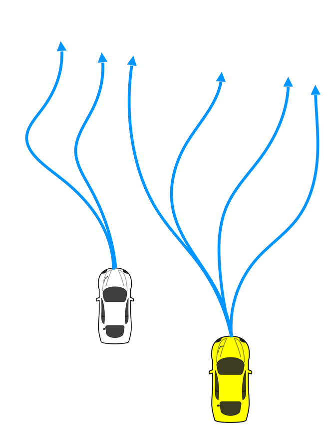
e.g. Boltzmann (Luce-Shepard) model

$$P(a_H \mid s, a_A) \propto e^{\beta Q_H(s, a_A, a_H)}$$

Our approach admits non-deterministic models of human decision-making.

Tactical Game: Trajectory Optimization

The autonomous vehicle plans its trajectory and predicts the humans' augmenting the short-horizon reward with the long-horizon strategic value.



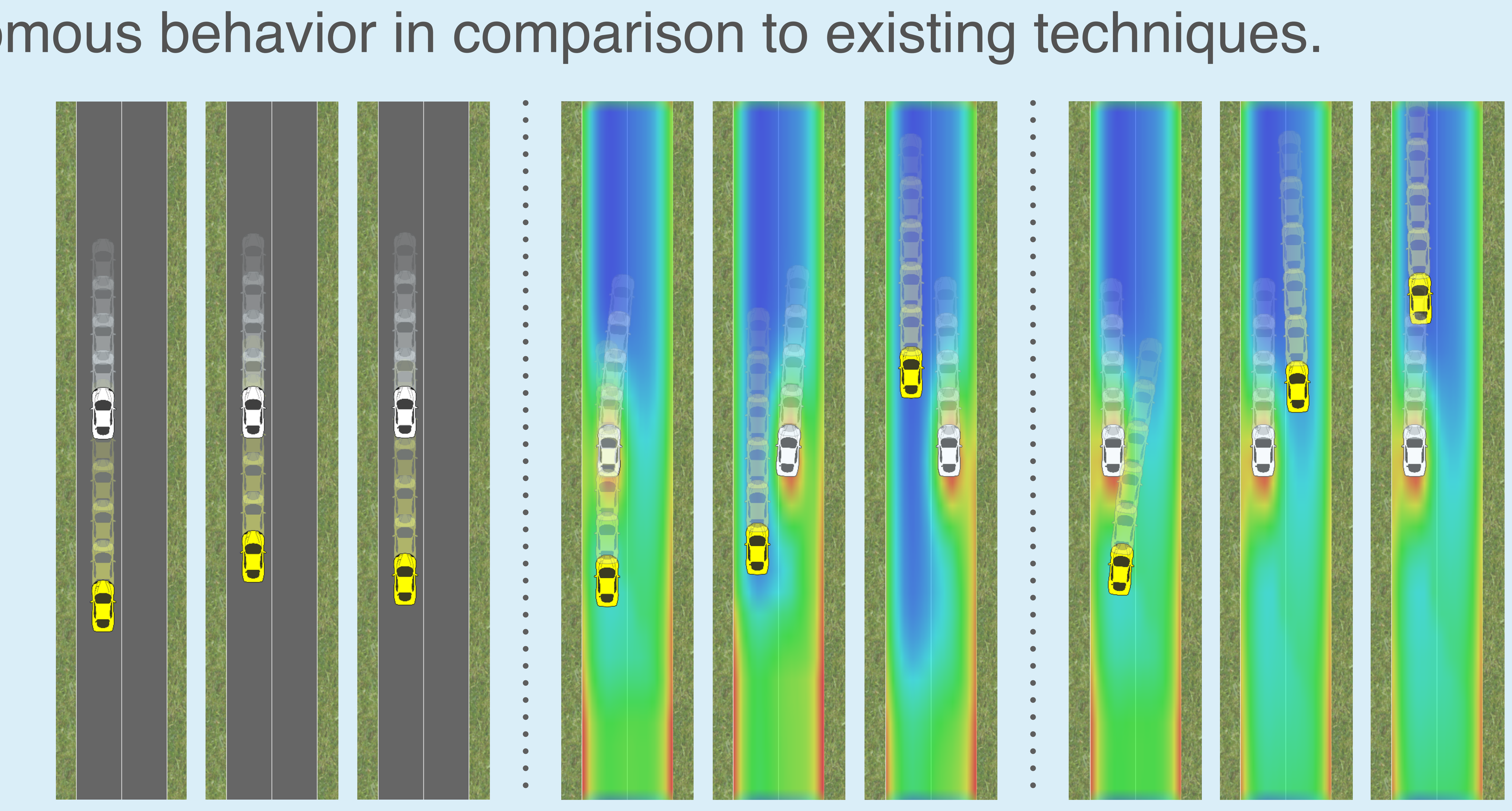
$$R_A(x^{0:M}, u_A^{0:M}, u_H^{0:M}) = \overbrace{\sum_{t=0}^{M} r_A(x^t, u_A^t, u_H^t)}^{\text{Tactical reward}} + \overbrace{V_A(g(x^t))}^{\text{Strategic value}}$$

$$R_H(x^{0:M}, u_A^{0:M}, u_H^{0:M}) = \sum_{t=0}^{M} r_H(x^t, u_A^t, u_H^t) + V_H(g(x^t))$$

Simulation Case Studies

Our case study results showcase richer, safer, and more effective autonomous behavior in comparison to existing techniques.

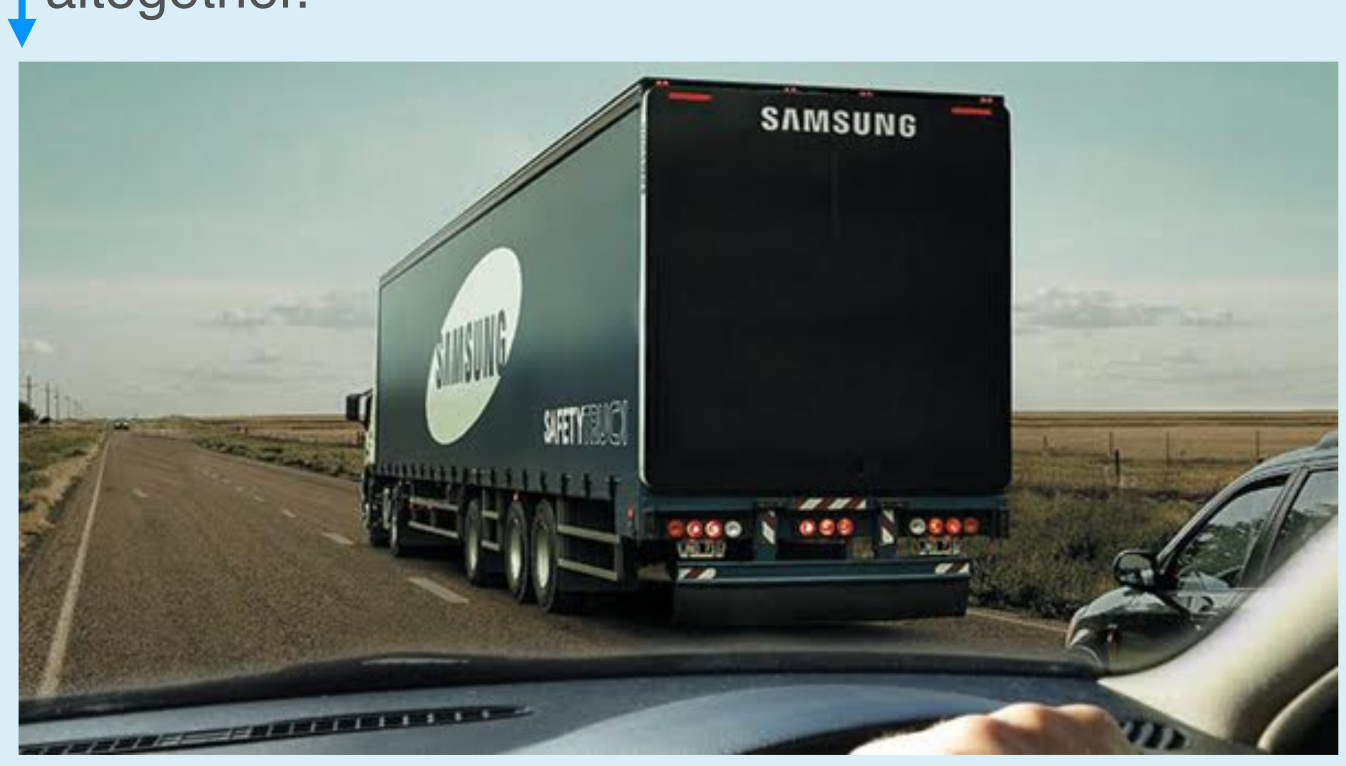
The hierarchical planner reliably completes the overtaking, seeking to influence the decisions of the human and adapting to the observed behavior.

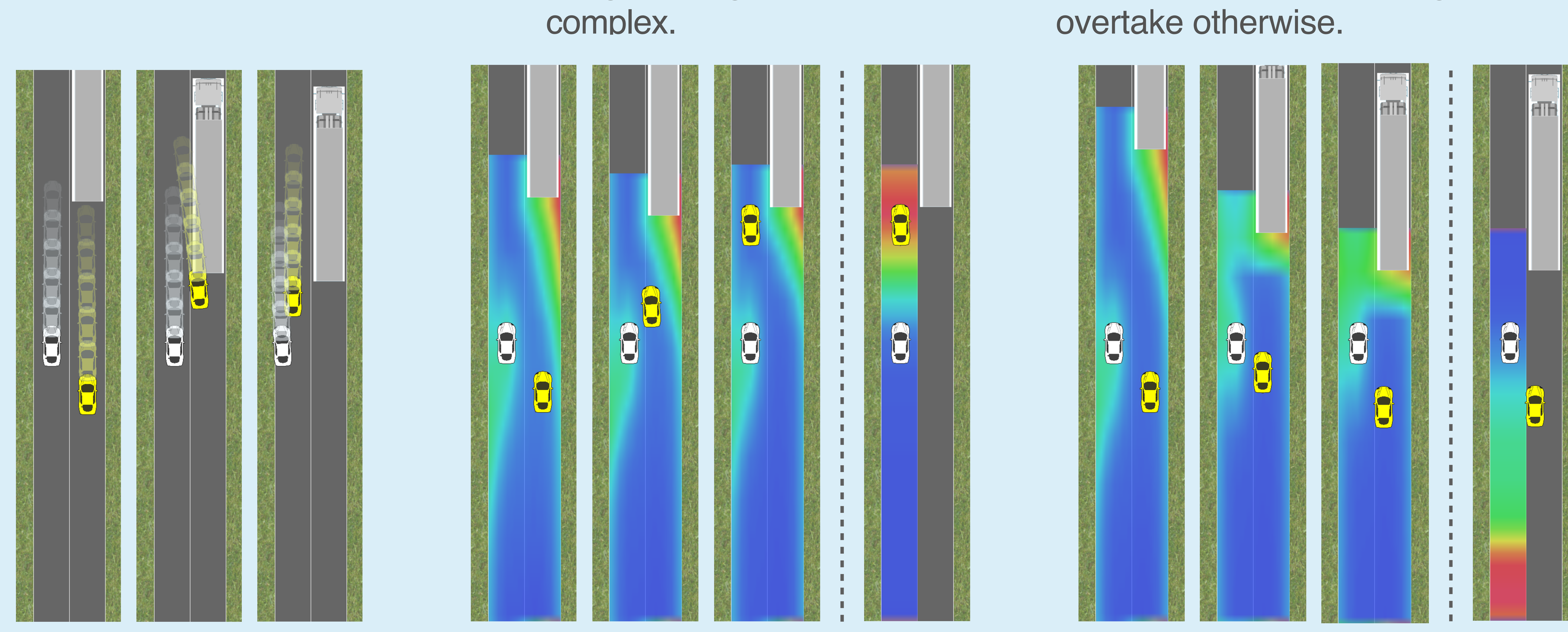


Low-level planner: overtaking too long/ complex.

Hierarchical planner: strategic value guides vehicle to incentivize human lane change, and overtake otherwise.

Another car tries to squeeze past and cut in before overtaking a truck, leaving two options: brake immediately and let the other car go or accelerate fast enough to dissuade the other driver altogether.





Low-level planner: dangerous cut-in emerges.

Hierarchical planner: depending on initial conditions, lead vehicle brakes to let the trailing vehicle cut in safely or accelerates to cut it off and force it to abort.