



HAL
open science

Approximate solutions to games of ordered preference

Pau de las Heras Molins, Eric Roy-Almonacid, Dong Ho Lee, Lasse Peters, David Fridovich-Keil, Georgios Bakirtzis

► To cite this version:

Pau de las Heras Molins, Eric Roy-Almonacid, Dong Ho Lee, Lasse Peters, David Fridovich-Keil, et al.. Approximate solutions to games of ordered preference. IEEE International Conference on Intelligent Transportation Systems (ITSC 2025), Nov 2025, Gold Coast, Australia. pp.3378-3385, <10.1109/ITSC60802.2025.11423775>. <hal-05565712>

HAL Id: hal-05565712

<https://hal.science/hal-05565712v1>

Submitted on 25 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ETALAB - Open licence

Approximate solutions to games of ordered preference

Pau de las Heras Molins*, Eric Roy-Almonacid*, Dong Ho Lee, Lasse Peters,
David Fridovich-Keil, and Georgios Bakirtzis

Abstract—Autonomous vehicles must balance ranked objectives, such as minimizing travel time, ensuring safety, and coordinating with traffic. Games of ordered preference effectively model these interactions but become computationally intractable as the time horizon, number of players, or number of preference levels increase. While receding horizon frameworks mitigate long-horizon intractability by solving sequential shorter games, often warm-started, they do not resolve the complexity growth inherent in existing methods for solving games of ordered preference. This paper introduces a solution strategy that avoids excessive complexity growth by approximating solutions using lexicographic iterated best response (IBR) in receding horizon, termed “lexicographic IBR over time.” Lexicographic IBR over time uses past information to accelerate convergence. We demonstrate through simulated traffic scenarios that lexicographic IBR over time efficiently computes approximate-optimal solutions for receding horizon games of ordered preference, converging towards generalized Nash equilibria.

I. INTRODUCTION

Complex agent decisions are often characterized by conflicting objectives. An autonomous vehicle, for example, must avoid collisions while also staying on the road, reaching a goal position, and obeying the speed limit. This problem is even harder when objectives of multiple agents conflict. Agents must determine which strategies are preferable, for example, going off-road to prevent a crash with another vehicle. Such decisions reflect an underlying comparative evaluation that ranks possible trajectories according to subjective preferences. We study the problem of comparative evaluations in transportation systems from the lens of *preference relations* that codify prioritized metrics; metrics that human drivers implicitly follow on the road.

A preference is a “total subjective comparative evaluation” [1]. In the context of an autonomous vehicle, a total subjective comparative evaluation is a rule that, given two trajectories, decides which of the two is preferred from the perspective of the controller. For multiagent systems too, preferences model the multiobjective and often conflicting requirements that the system must adhere to for producing good behavior as subjective comparative evaluations (Fig. 1).

One way to express the tension between prioritized metrics of multiple agents is to express them as lexicographic optimization problems in a game of ordered preference [4]. Games of ordered preference are multi-player games with prohibitive

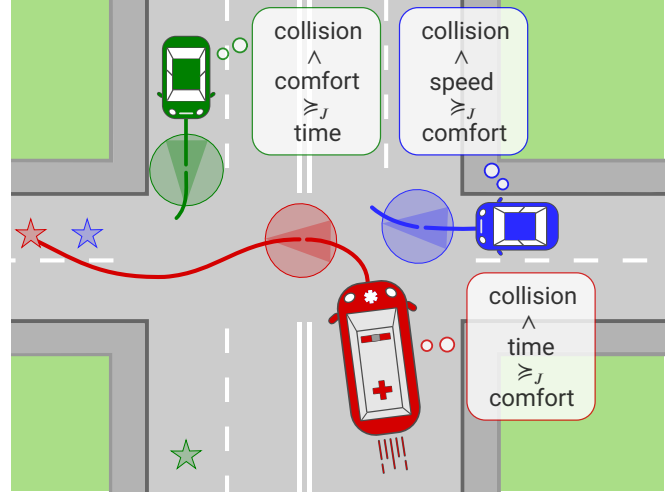


Fig. 1. Agents operate under shared constraints, but are incentivized by different preferences (adapted from Zanardi et al. [2], Mendes Filho et al. [3], and Lee et al. [4]). On the road, situations arise where a vehicle has to coordinate with other vehicles and must relax a low-priority metric to preserve more important metrics, such as giving way to an ambulance.

computational cost. Solving these games with a receding horizon alleviates this cost through shorter time horizons and the introduction of feedback. However, the receding horizon setting does not address the rapid dimensional increase caused by adding players or preference levels. Iterated best response (IBR) efficiently solves multi-player games by decomposing them into single-player problems, where each player optimizes assuming fixed opponent strategies. IBR convergence can be slow due to the need to explore large opponent decision spaces. We propose “lexicographic IBR over time,” an extension of IBR that incorporates past information, available in receding horizon, to make predictions about future opponent decisions that accelerate convergence to generalized Nash equilibria for games of ordered preference.

To study this area, this paper asks:

(?) *Can we efficiently compute approximate-optimal trajectories for games of ordered preference?*

To make progress on this question, we

- 1) find that lexicographic IBR over time approximates optimal solutions for games of ordered preference; and
- 2) design experiments demonstrating that games of ordered preference are an effective tool for analyzing interactive preferences in transportation systems.

*Equal contribution and corresponding authors.

P. de las Heras Molins (pau.de.las.heras@estudiantat.upc.edu) and E. Roy-Almonacid (eric.roy@upc.edu) are with Universitat Politècnica de Catalunya. D. H. Lee and D. Fridovich-Keil are with The University of Texas at Austin. L. Peters is with Technische Universiteit Delft. G. Bakirtzis is with LTCI, Télécom Paris, Institut Polytechnique de Paris.

A. Related work

We have been primarily motivated to study games of ordered preference by Zanardi et al. [2] and Lee et al. [4]. We explicitly build on this research to compute for a receding horizon, rather than a finite-time horizon, to improve the scalability of these games. We have additionally found insights in receding horizon games [5, 6], although the formulation is not directly applicable to the problem of preference ordering.

Prior work, closely related to the formulation of games of ordered preference in this paper, is on *prioritized metrics* and *game-theoretic planning*.

PRIORITIZED METRICS Posetal games [2] formalize games with prioritized metrics through partial orders over system designs [7], using symbolic structures to encode constraint and preference hierarchies [8]. While posetal games focus on discrete decision spaces, we adopt lexicographic optimization [9], that extends prioritized metrics to continuous action spaces [4]. However, the lexicographic nested problem structure introduces computational challenges: lower-priority objectives are optimized only after higher-priority ones are satisfied, creating a hierarchy of interdependent subproblems. To mitigate this computational intractability, prior work transcribes the lexicographic hierarchy into mathematical programs with complementarity constraints [10], with the addition of relaxation schemes [4]. Yet, coupled constraints cause scalability to remain limited. By framing the problem within a receding horizon setting, we exploit the temporal decoupling inherent to IBR algorithms, which are known to be computationally tractable.

GAME-THEORETIC PLANNING Game-theoretic traffic management is needed to analyze the interactive nature of preferences between autonomous agents [11]. Efficient solution strategies exist for game-theoretic planning [12–15], but solving games that involve nested optimization problems typically requires approximate solutions, similar to those used in solving trajectory games with continuous action spaces [16–18]. Moreover, predictions of other agents’ intentions combined with planning increase the accuracy of approximate solutions [19, 20]. Receding horizon in game-theoretic planning has been studied in the classical formulation [21] and in variants—where objects are divided into immediate and mid-term metrics [22]. These approaches do not consider the problem of preferences in agent interactions.

II. PRELIMINARIES

In this section, we review key concepts in trajectory games, lexicographic minimization, and games of ordered preference.

NOTATION For any non-negative number n , we define the set of integers $[n] := \{1, 2, \dots, n\}$. We use boldface to represent a time-indexed vector with length of a finite-time horizon game $\mathbf{z} := \{z_0, z_1, \dots, z_{T_g-1}\}$. We time index this vector between $t_0 \in \mathbb{N}$ and $t_1 \in \mathbb{N}$ as the vector $\mathbf{z}_{t_0:t_1} := \{z_{t_0}, z_{t_0+1}, \dots, z_{t_1}\}$. We use superscripts as in \square^i to denote the agent $i \in [N]$. Negation is used to include all agents but one, $\square^{-i} := \square \setminus \square^i$.

A. Trajectory games

Trajectory games are non-cooperative, multiagent, general-sum constrained dynamic games. Solutions to trajectory games amount to finding generalized Nash equilibrium (GNE) in which all agents adopt optimal discrete-time trajectories from which none can unilaterally deviate without incurring an increase in cost. In a trajectory game the state and action spaces for each agent $i \in [N]$ are defined as $\mathcal{X}^i \subseteq \mathbb{R}^n$ and $\mathcal{U}^i \subseteq \mathbb{R}^m$ respectively, with states usually encoding the position of agents in the environment and actions encoding their control inputs. We assume that states and actions can be split by agent. At discrete-time step t , the game decision of each agent, $\mathbf{z}_t^i = [x_t^i, u_t^i]$ is composed by the i th agents’ current state, $x_t^i \in \mathcal{X}^i$, and control input to be applied, $u_t^i \in \mathcal{U}^i$, with the resulting decision space being $\mathcal{Z}^i \subseteq \mathcal{X}^i \times \mathcal{U}^i$. We define the trajectory of each agent as a sequence of decisions $\mathbf{z}^i = \{z_0^i, \dots, z_{T_g-1}^i\}$ over the time horizon of the game T_g .

An individual cost function $J^i : \mathcal{Z}^i \times \mathcal{Z}^{-i} \rightarrow \mathbb{R}$ encodes agent objectives over a trajectory, as well as a set of private equality, g^i , and inequality, h^i , constraints, with $g^i : \mathcal{Z}^i \rightarrow \mathbb{R}$ and $h^i : \mathcal{Z}^i \rightarrow \mathbb{R}$. Agents must also satisfy a set of shared equality, g^s , and inequality, h^s constraints, that consider the global system state and for which all agents are equally responsible, with $g^s : \mathcal{Z}^i \times \mathcal{Z}^{-i} \rightarrow \mathbb{R}$ and $h^s : \mathcal{Z}^i \times \mathcal{Z}^{-i} \rightarrow \mathbb{R}$. Solutions consist of joint trajectories for all the agents, \mathbf{z} .

We can package the above as a generalized Nash equilibrium problem (GNEP) [18]. For each agent i , the open loop, complete information GNEP is defined as

$$\begin{aligned}
 & \underset{\mathbf{z}^i}{\text{minimize}} && J^i(\mathbf{z}^i, \mathbf{z}^{-i}; \theta^i) && \text{dynamics} \\
 & \text{subject to} && x_{t+1}^i = f^i(x_t^i, u_t^i) && \text{for all } t \in [T_g - 1] \\
 & && && \text{initial state} \\
 & && x_1^i = \hat{x}_1^i && \\
 & && g^i(\mathbf{z}) = 0 && \\
 & && h^i(\mathbf{z}) \geq 0 && \\
 & && g^s(\mathbf{z}, \mathbf{z}^{-i}) = 0 && \\
 & && h^s(\mathbf{z}, \mathbf{z}^{-i}) \geq 0. && \\
 & && && \text{shared constraints} \\
 & && && \text{e.g., collision avoidance}
 \end{aligned} \tag{1}$$

The objective function, additionally parametrized by θ^i , and constraints take as arguments the trajectory \mathbf{z} of all agents, differentiating the i th agent trajectory \mathbf{z}^i from the others \mathbf{z}^{-i} .

B. Lexicographic minimization

For each $k \in [K]$ let $J_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be the k th objective function of the decision variables residing within a feasible set \mathcal{Z} . We define a total order \succ_J on decision variables as follows: for any $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$, we say that $\mathbf{z} \succ_J \mathbf{z}'$ if and only if $\mathbf{z} = \mathbf{z}'$ or $J_k(\mathbf{z}) < J_k(\mathbf{z}')$ for the smallest k such that $J_k(\mathbf{z}) \neq J_k(\mathbf{z}')$. In other words, J_1 holds the highest priority and J_K the lowest.

A lexicographic minimum of a feasible set $\mathcal{Z} \in \mathbb{R}^K$ is a decision variable $\mathbf{z}^* \in \mathcal{Z}$ for which $\mathbf{z} \preceq_J \mathbf{z}^*$ for all $\mathbf{z} \in \mathcal{Z}$.

Computing all lexicographic minima amounts to a nested optimization problem. The total order \succcurlyeq_J encodes a strict hierarchy, meaning that given any two decision variables $\mathbf{z}_h, \mathbf{z}_l$, with $h < l$, there exists a clear preference for one over the other, $\mathbf{z}_h \succcurlyeq_J \mathbf{z}_l$ (we say \mathbf{z}_h has higher priority than \mathbf{z}_l), and no two priorities are at the same level of importance.

C. Games of ordered preference

In a game of ordered preference agents pursue multiple, hierarchically ranked objectives. Unlike other games, where agents trade off objectives, here agents resolve conflicts through lexicographic minimization, strictly prioritizing goals.

At each preference level $k \in [K^i]$, agent i selects a trajectory \mathbf{z}_k^i from a feasible set that is constrained by the lower (more prioritized) preference levels. Agents know the final trajectories of the other agents, \mathbf{z}^{-i} , when optimizing at level k . The game of ordered preference is equivalent to a lexicographic minimization problem that computes a GNE, where no agent can improve a higher-priority objective without violating constraints or compromising lower-priority goals. The solution to a game of ordered preference is found by jointly solving the following lexicographic minimization problem for each agent i across the K^i priority levels,

$$\begin{aligned}
& \underset{\mathbf{z}_{K^i}^i}{\text{minimize}} && J_{K^i}^i(\mathbf{z}_{K^i}^i, \mathbf{z}^{-i}; \theta^i) \\
& \text{subject to} && \mathbf{z}_{K^i}^i \in \underset{\mathbf{z}_{K^{i-1}}^i}{\text{argmin}} J_{K^{i-1}}^i(\mathbf{z}_{K^{i-1}}^i, \mathbf{z}^{-i}; \theta^i) \\
& && \vdots \quad \begin{array}{l} \text{principal preference} \\ \text{e.g., speed limit} \end{array} \\
& \text{subject to} && \mathbf{z}_2^i \in \underset{\mathbf{z}_1^i}{\text{argmin}} J_1^i(\mathbf{z}_1^i, \mathbf{z}^{-i}; \theta^i) \quad (2) \\
& \text{subject to} && \mathbf{z}_1^i \in \mathbb{R}^n \quad \begin{array}{l} \text{private constraints} \\ \text{i.e., feasibility region} \end{array} \\
& \text{such that} && g^i(\mathbf{z}^i, \mathbf{z}^{-i}) = 0, \text{ and} \\
& && h^i(\mathbf{z}^i, \mathbf{z}^{-i}) \geq 0 \\
& \text{such that} && g^s(\mathbf{z}) = 0, \text{ and} \\
& && h^s(\mathbf{z}) \geq 0. \\
& && \begin{array}{l} \text{shared constraints} \\ \text{e.g., collision avoidance} \end{array}
\end{aligned}$$

The optimal solution for each agent, that is, the actual strategy that they will want to follow, is the solution computed for the first level of the lexicographic minimization problem, $\mathbf{z}_{K^i}^i$ in Eq. (2), which embeds all the feasible set defining constraints of inner levels. For notational simplicity, we refer to the solution of the game for each agent i as $\mathbf{z}^i := \mathbf{z}_{K^i}^i$.

In a trajectory game, the set of private constraints of an agent that restrict their original feasibility regions typically include their dynamics and initial state. The private constraints need to take into account the preference structure of the other agents, making $g^i : \mathcal{Z}^i \rightarrow \mathbb{R}$ and $h^i : \mathcal{Z}^i \rightarrow \mathbb{R}$. Additionally, a common shared constraint for these type of games is collision avoidance (Section II-A). This enables a more expressive definition of an agent's objectives, including those that have

utmost importance (such as safety constraints or following the speed limit) and should only ever be suboptimal in the presence of a hard constraint. For example, an autonomous vehicle should always prioritize not breaking the speed limit (higher priority objective) even at the cost of reaching their goal (lower priority objective) at a later time, only ever speeding past it to avoid a collision (hard constraint).

It is often desirable to flatten the hierarchy of nested subproblems in Eq. (2) into a single level. For each agent, beginning with their innermost problem, we derive the Karush-Kuhn-Tucker (KKT) conditions and incorporate the resulting dual variables as induced primals in the outer problems. This effectively constrains the feasible set, ensuring that solutions conform to the lexicographic order of preferences. The resulting formulation is a mathematical program with complementarity constraints for each agent, which can be regularized by a relaxation scheme that expands the feasible set to solve the mixed complementarity problem (MCP) [4].

III. EFFICIENTLY FINDING SOLUTIONS TO GAMES OF ORDERED PREFERENCE

In this section, we formulate games of ordered preference into a receding horizon setting and produce algorithms that exploit lexicographic IBR over time for efficient computation.

A. Receding horizon games of ordered preference

Games of ordered preference introduce considerable complexity due to the nested structure of the optimization problem they generate. Even when the hierarchy is reduced to single-level problems, the dimensionality of the decision variables increases substantially. Solving trajectory games framed in terms of dynamical systems within the context of games of ordered preference is even more challenging due to the higher dimensionality needed to represent the dynamics, along with spatiotemporal constraints and time-indexed objectives, particularly when considering long, fixed-time horizons (Section II-A). The challenge of high dimensionality restricts their use to instances with short, fixed-time horizons. This limitation in turn diminishes agents' ability to adapt to medium- and long-term environmental changes.

A receding horizon game of ordered preference mitigates this problem by partitioning the solution space into manageable segments (Fig. 2). In a receding horizon game with length of a finite-time horizon, T_g , agents periodically re-optimize their strategies. At each decision-making stage, however, they focus on a reduced time horizon, $T \ll T_g$, repeatedly solving optimization subproblems with the latest state observations serving as the initial conditions [23]. At each decision-making stage, the agents extract and execute only the first T_l decisions of the newly computed optimal joint strategy, where $0 < T_l \leq T$ is called the *turn length* [24]. These decision-making stages occur every T_l time steps, continuing until the end of the game at time T_g .

A receding horizon variation of a game of ordered preference involves each rational agent $i \in [N]$ finding an optimal strategy that lexicographically minimizes their ordered cost functions J^i over the game time horizon $t \in \{0, \dots, T_g - 1\}$.

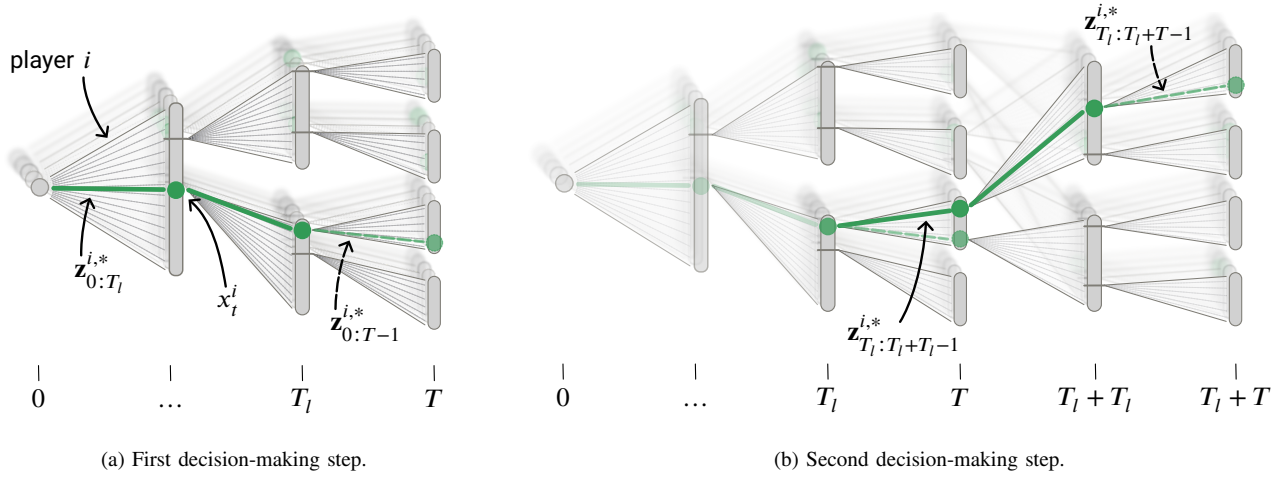


Fig. 2. In a trajectory game, each agent’s strategy defines an optimal trajectory governed by a continuous state that evolves through actions selected from a continuous control input space at discrete time intervals. Agents compute a receding horizon to dynamically adapt their trajectories in response to environmental changes, periodically updating their strategies every T_l time steps. At each decision stage, agents solve a game-theoretic optimization problem over a finite time horizon T , generating optimal trajectories (dashed lines). From these trajectories, agents execute only the first T_l control inputs (solid lines) before resolving the game with updated environmental information.

The optimal strategy along the entire game horizon T_g , which we denote as $\mathbf{z}^{i,*}$, is incrementally built by aggregating the executed steps of the partial solutions obtained at each decision-making stage.

At each decision-making stage, occurring every T_l time steps, a game of ordered preferences is solved over a reduced time horizon T . Each agent minimizes the partial cost function

$$\ell^i(\mathbf{z}_{t:t+T-1}^i, \mathbf{z}_{t:t+T-1}^{-i}; \theta^i).$$

The partial cost ℓ^i for each agent i , with $\ell^i: \mathcal{Z}^i \rightarrow \mathbb{R}$, is the cost of executing the next T decisions $\mathbf{z}_{t:t+T-1}^i$ given the other agent’s decisions $\mathbf{z}_{t:t+T-1}^{-i}$. It differs from the objective function J^i in Eq. (2) in that the latter accounts for the entire game horizon T_g , while ℓ^i gives the cost for a single reduced horizon solution for agent i computed at time step t ,

$$J^i(\mathbf{z}^i, \mathbf{z}^{-i}; \theta^i) = \sum_{t=0}^{T_g-1} \ell^i(\mathbf{z}_{t:t+T-1}^i, \mathbf{z}_{t:t+T-1}^{-i}; \theta^i). \quad (3)$$

We repeat Eq. (3) at times $t = m \cdot T_l$ for all positive integers m such that $m \cdot T_l < T_g$.

The solution of a single decision-making stage at time step t is a joint optimal strategy $\mathbf{z}_{t:t+T-1}^*$. Agents computing a receding horizon will find at once their optimal strategy for the next T time steps,

$$\mathbf{z}_{t:t+T-1}^{i,*} = \underset{\mathbf{z}_{t:t+T-1}^i}{\operatorname{argmin}} \ell^i(\mathbf{z}_{t:t+T-1}^i, \mathbf{z}_{t:t+T-1}^{-i,*}; \theta^i).$$

The optimal solution for each agent i at time step t corresponds to the solution of the lowest priority level in Eq. (2). After each decision-making stage, only the first T_l decisions of the solution for all agents, $\mathbf{z}_{t:t+T_l-1}^*$, are aggregated into the final solution $\mathbf{z}_{0:T_g-1}^*$, with the associated control inputs executed by the agents to evolve their states. At step $t + T_l$, a new decision-making stage takes place, which solves the game again for $t + T_l + T$, using initial

state $\mathbf{x}_{t+T_l}^i$. This computation is repeated until the end of the game is reached at time T_g . The algorithmic interpretation of this system involves a measurement step at time t where the global state \mathbf{x}_t is observed after evolving the system for T_l time steps following the previous decision-making stage (Algorithm 1).

When comparing receding horizon formulations of games against the corresponding fixed-time horizon ones, the benefits are two fold. First, by solving for only $T \ll T_g$ time steps at a time, a receding horizon approach circumvents the computational complexity associated with accounting for the entire solution at once. Second, deferring the resolution of the final steps allows for adaptive decision making in response to evolving environmental conditions, allowing agents to reconsider their strategies over time (Fig. 2).

B. Lexicographic IBR over time

Games of ordered preference incur a significant increase in the problem’s dimensions for each preference level. The significant increase arises from the accumulation of extra equality and inequality constraints and induced primals derived from the KKT conditions at each level during the flattening of the nested optimization problem [4]. The flattening of multiple preference levels, while necessary to solve games of ordered preference, can quickly grow the problem dimensions to become computationally intractable. One way to mitigate the problem of high dimensionality is to partition the coupled multi-player game of ordered preference into smaller single-player subproblems which can be solved with extensively used algorithms such as IBR.

In the standard formulation of IBR, multiple agents repeatedly take turns to compute their best response considering other players’ strategies. Under certain assumption that are true for games of ordered preference IBR converges to a Nash equilibrium [25]. IBR in the receding horizon setting uses information about past decisions of other agents to better

DATA: MCP for the game of ordered preference,
game horizon T_g , receding horizon T ,
turn length T_l , and initial states \mathbf{x}_0

RESULT: Optimal joint strategies \mathbf{z}^*

Start at $t = 0$ with initial states \mathbf{x}_0 for each agent i

WHILE $t < T_g$ DO

 Measure global state \mathbf{x}_t at time t

 Solve MCP at \mathbf{x}_t with time horizon T

 Evolve state for time T_l using $\mathbf{z}_{t:t+T_l-1}^*$

 Recede horizon: $t \leftarrow t + T_l$

$\mathbf{z}^* \leftarrow \mathbf{z}_{0:T_g-1}^{i,*}$ for each agent i

Algorithm 1. Receding horizon for games of ordered preference.

predict their future best responses, termed ‘‘IBR over time.’’

In lexicographic IBR over time, the multiagent problem is split in N single-player games, with each player performing lexicographic optimization of their preferences while considering other players’ strategies as fixed. The aggregate solution of the single-player games is an approximation of the GNE described in the multiagent formulation. Lexicographic IBR over time differs from IBR by starting with a more accurate initial guess of the trajectories of other agents and stopping at a fixed number of iterations. In particular, IBR over time improves standard IBR efficiency by warmstarting agents’ best responses, which reduces the number of iterations needed to converge to an equilibrium.

Approximate solutions are acceptable when they are close to the optimum. Lexicographic IBR over time produces approximate-optimal solutions, provided there are good predictions of other agents’ strategies and an acceptable number of iterations. An effective strategy for generating good predictions is to use the trajectories of other agents computed in the previous decision-making stage. Further refinement occurs in additional IBR iterations, where each agent sequentially considers the running best responses of others until convergence to the GNE is achieved or a maximum number of iterations (L in Algorithm 2) is reached.

The complete predictions for the first iteration of IBR at a decision-making stage are obtained by evolving the current measured state using the dynamics with the action sequences,

$$z_{t+1}^i = [x_{t+1}^i, u_{t+1}^i] \approx \left[\overset{\text{preceding action}}{f^i(x_t^i, u_t^i)}, \overset{\text{dynamics}}{u_t^i} \right].$$

The action sequences are obtained by shifting those of previously computed trajectories and padding them with additional actions to fill the time horizon T (Fig. 3).

In IBR, the lexicographic minimization of each player’s objectives can be achieved by successively solving single-player optimization problems. For each each preference level k of player i , we solve a minimization problem with cost function ℓ_k^i as its objective. In this setting, the other

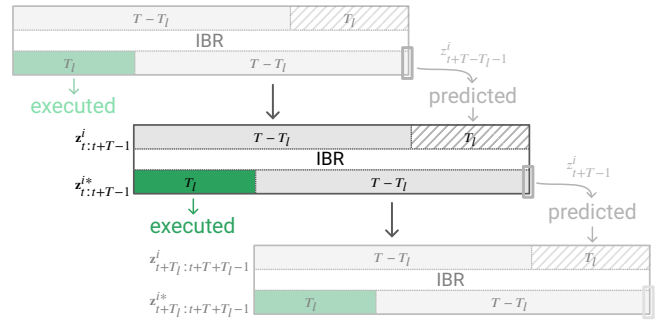


Fig. 3. When solving using IBR, the predictions of the other agents’ trajectories are based on the solutions obtained at the last decision-making stage. The previous solutions are shifted in time and then padded with additional action, which can be a null action u_0^i or the same action as the previous step u_{t-1}^i .

player’s trajectories become another parameter of the game, included in θ^i . The optimal cost value computed at any level $y_k^{i,*} = \ell_k^i(\mathbf{z}_k^i; \theta^i)$ is recorded. Then, in each subsequent outer level, we add additional inequality constraints to preserve the optimality with respect to previously optimized inner objectives [26]. For any level k , the added constraints are

$$\ell_j^i(\mathbf{z}_k^i; \theta^i) \leq y_j^{i,*}$$

for each previous level $1 \leq j < k$. Given that the optimization method used is not optimal since it is based on iterated approximations, by using inequality constraints we ensure that if in subsequent optimizations better solutions for previous levels’ objectives are encountered they won’t be rejected.

The IBR over time solution, described in Algorithm 2, acts as a replacement of the MCP solving step in Algorithm 1 in the single-player formulation.

In contrast to the single, coupled optimization problem solved for the game of ordered preference, IBR over time solves a total of $\sum_{i=1}^N |K^i|$ problems at each decision-making stage, where K^i represents the number of preference levels for agent i . However, the reduced dimensionality of the individual problems results in significantly faster computation times. Moreover, the IBR over time solution can be computed in parallel for each agent i , rather than sequentially. This parallelization can further enhance processing efficiency, particularly in games with many agents.

IV. EXPERIMENTS

In this section we evaluate the performance of IBR over time for solving games of ordered preference. We design experiments to (1) examine how different preference relations yield qualitatively different solutions for the same scenario and (2) quantitatively compare the efficiency of IBR over time against the standard formulation of games of ordered preference in receding horizon.

A. Implementation details

We implement the code in Julia [27], build upon the code of Lee et al. [4], and make extensive use of TrajectoryGames-Base.jl [24] to instantiate and solve games of ordered preference in the receding horizon setting. The benchmarks use the

DATA: MCP for the game of ordered preference,
current time t , preceding trajectories
 $\mathbf{z}_{t:t+T-T_i-1}^i$, time horizon T , maximum
iterations L , and convergence tolerance ϵ
RESULT: Approximate-optimal joint strategies $\mathbf{z}_{t:t+T-1}^*$

$\mathbf{z}_{0:T-1} \leftarrow$ Shift and pad preceding trajectories
Best responses $\leftarrow \mathbf{z}_{0:T-1}$
FOR 1 TO L DO
 FOR $i = 1$ TO N DO
 Parameters $\theta^i \leftarrow$ best response of agents $-i$
 $\mathbf{z}_{0:T-1}^i \leftarrow \operatorname{argmin}_{\mathbf{z}_{0:T-1}^i} \ell^i(\mathbf{z}_{0:T-1}; \theta^i)$
 Update best response of agent i
 IF *solution improvement* $< \epsilon$ THEN
 BREAK

Algorithm 2. A single IBR solution in games of ordered preference.

package ParametricMCPs.jl [28], which is a wrapper around the PATH solver [29]. We package the code for running the following experiments as LEXIBROverTime.jl [30].

All benchmarks are run on a single core of a machine with no additional load.

B. Qualitative evaluation of approximate solutions

EVALUATION SCENARIO To qualitatively assess how different preference relations influence the solution of a game of ordered preference, given identical parameters, we consider a road navigation problem involving three players: two cars driving in opposite lanes of a road and an ambulance attempting to rush past them in response to an emergency. We present two simulation instances with the same parameters (road layout, initial state, dynamics, constraints, etc.), but with different preference relations (Fig. 4). The game is solved using IBR over time with a single IBR iteration ($L = 1$) and the same time horizon, turn length, and simulation times for both scenarios.

MAIN RESULT ① – IBR OVER TIME DISTINGUISHES DIFFERENT PREFERENCE RELATIONS Trajectories of agents with different preference relations vary (Fig. 4). In the highway scenario, the green car slightly veers out of its lane to allow the ambulance to pass, while the blue car temporarily leaves the road to avoid a potential collision. In contrast, in the urban scenario, the green car remains at the edge of its lane, and the blue car executes a sharp turn towards the center of the road to avoid pedestrians, while the ambulance overtakes in the opposite lane. This demonstrates how IBR over time identifies solutions that respect the preference hierarchy of the agents, thereby replicating the outcomes derived from the standard formulation.

C. Efficiency of lexicographic IBR over time

EVALUATION SCENARIO To assess the efficiency of IBR over time we consider a simplified variant of the previous problem: two cars driving along parallel lanes of a road

TABLE I

PERFORMANCE COMPARISON BETWEEN BASELINE AND IBR OVER TIME WITH DIFFERENT ITERATION LIMITS L AND NUMBER OF PREFERENCE LEVELS K . TIMES AND L_1 DISTANCES ARE AVERAGED ACROSS RUNS.

Method	$K = 2$		$K = 3$	
	t_{solve}	L_1 distance	t_{solve}	L_1 distance
Baseline	44.51s	–	2676s	–
$L = 1$	0.36s	3.28×10^{-4}	0.32s	5.45×10^{-4}
$L = 2$	0.59s	1.22×10^{-5}	0.63s	6.56×10^{-5}
$L = 3$	0.85s	4.64×10^{-6}	0.75s	5.66×10^{-5}
$L = 5$	1.39s	1.78×10^{-6}	1.20s	4.96×10^{-5}
$L = 10$	2.67s	1.71×10^{-6}	2.32s	3.03×10^{-5}
$L = 20$	5.40s	1.28×10^{-6}	4.55s	8.51×10^{-5}
$L = 1000$	269.4s	2.45×10^{-6}	273.27s	4.14×10^{-5}

in the same direction, with an ambulance attempting to overtake them. The fact that all vehicles are traveling in the same direction reduces the risk of collision, resulting in less restrictive optimization problems that can typically be solved using both IBR over time and the baseline method. We conduct a Monte Carlo study with 20 variations of the scenario, generated by adding random permutations sampled from a uniform distribution centered around the initial states (position and speed) of the agents. The experiments are run with two different numbers of preference levels K , common to all players, in order to account for the increased complexity of the optimization problems associated with a higher number of priority levels.

BASELINE For the baseline, we use an adaptation of the solver for games of ordered preference in their standard formulation [4], which operates in the receding horizon setting. We use the same hyperparameters in both methods, including the relaxation iterations specified by Lee et al. [4].

EVALUATION METRICS The primary evaluation metrics for performance comparison are the average time required to compile the problems t_{compile} and the average time taken to solve each of them t_{solve} . The compile time is a one-time overhead not included in the solution time. For comparison, the solution times for IBR over time are expressed as a percentage of the baseline solution time. Since there might be a continuum of equilibria, that is, multiple optimal solutions, we will evaluate the quality of the solutions by computing the L_1 distance between the solution trajectory and what would be the solution at the next IBR iteration $L + 1$.

MAIN RESULT ② – IBR OVER TIME EFFICIENTLY SOLVES GAMES OF ORDERED PREFERENCE The results demonstrate that IBR outperforms the baseline in both problem compilation and solution times, with a substantial margin, even as the number of iterations L increases (Table I). The differences in solution times become more pronounced as the number of preference levels increases, with the baseline showing a much larger degradation in performance. The average solution time for IBR is directly proportional to the number of IBR iterations. A maximum number of iterations can be adjusted

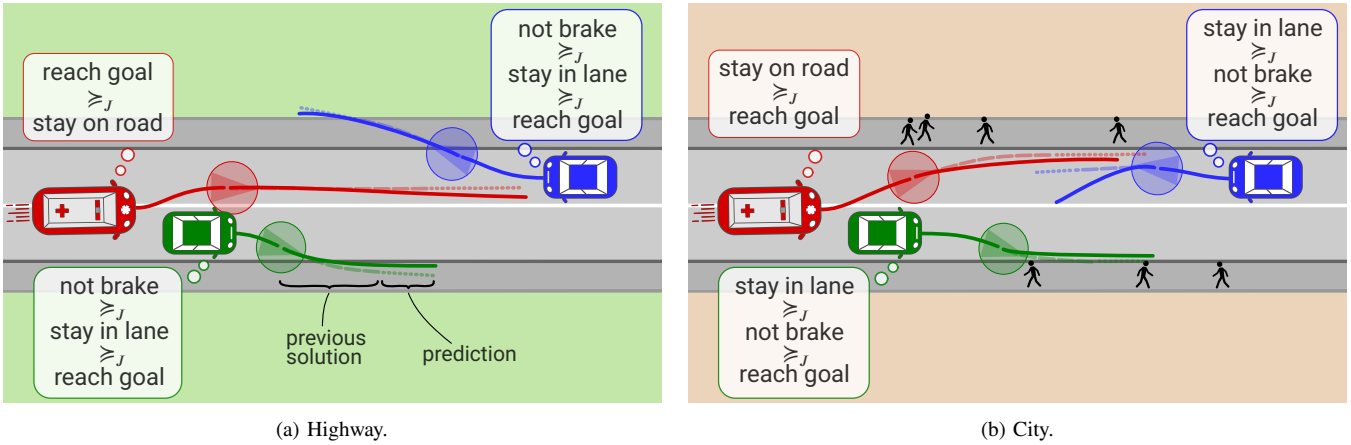


Fig. 4. Different preference relations model varying contexts, resulting in qualitatively different agent behaviors. In a highway scenario, vehicles prefer leaving their lane to avoid braking hard because there are no pedestrians, while in an urban scenario they prefer to stay within the road to not risk hitting pedestrians. Dashed lines show each agent’s trajectory as predicted by other agents. Solid lines represent their actual computed trajectories.

to balance the optimality of solutions with computational constraints. The problem compilation time is not included in the results, summarized in Table I, given that its value does not depend on the number of iterations L . The average baseline compilation times t_{compile} are 1.50s and 9.37s for preference levels $K = 2$ and $K = 3$ respectively. For the same preference levels, IBR over time produces average compilation times that are significantly lower, 0.75s and 1.10s, primarily due to the reduced size of the problems and increased solver efficiency. The compilation of all subproblems in IBR over time takes less total time than the compilation of a single problem in the baseline. Shorter compilation times enable faster transitions between preference relations, which are embedded in the problem’s structure. Additionally, the L_1 distance of the solution at iteration L with respect to the next iteration $L + 1$ decreases rapidly within a few iterations, indicating that the trajectories stabilize and become approximate-optimal. The final average L_1 distance does not reach zero (Table I). This phenomenon arises because IBR does not converge to a stable equilibrium for some problems. The limit, L , ensures the algorithm terminates in those cases.

D. Discussion

The experimental results demonstrate that it is possible to efficiently find approximate solutions to games of ordered preference (answering question (?)). Lexicographic IBR over time enables tractable computation of single-player versions of games of ordered preference while still respecting the agents’ preference hierarchy. Although the solutions are not guaranteed to be optimal compared to the standard formulation, the number of IBR iterations can be adjusted to reduce the gap to the optimal solution, provided time and computational resources are not constrained. However, practical results indicate that even with a single IBR iteration, using previously computed trajectories as a base for predicting the other agents’ trajectories provides a reasonable approximation.

The benchmarks demonstrate that IBR over time reduces solution times for games of ordered preference compared to existing methods, especially as the number of preference

levels increases. This improved computational tractability enables solving larger problems with more preferences and longer time horizons—both crucial factors for developing agents capable of complex behaviors and effective adaptation to short- and long-term environmental changes. However, the lexicographic structure of games of ordered preference can lead to dominant strategies, where agents primarily optimize for higher-priority objectives. This tendency toward dominant strategies can be mitigated through careful modeling of agents’ preferences, ensuring that the feasible regions defined by successive preference levels still allow for exploring lower-priority objectives. The results also show that the solution stabilizes, that is, the solution changes less with respect to the previous iteration $L - 1$, as number of iterations L grows. This stability of solutions means that with just a few iterations we can achieve approximated-optimal solutions.

V. CONCLUSION

Autonomous vehicles are objective decision makers, yet the quality of those decisions must be measured against the subjective, context-aware choices of human drivers. Preference relations model prioritized metrics in multi-objective optimization, which capture the subjective decision making of human drivers. Accounting for the context-varying preferences and intentions of all vehicles in the road is a necessary condition for intelligent traffic management. Games of ordered preference have demonstrated compelling support for addressing this requirement. However, the computational cost of considering preference hierarchies has limited their inclusion in real-time systems’ design. Beyond real-time trajectory optimization, autonomous navigation also demands that agents are able to react to the changing environments. Our proposed algorithm, lexicographic IBR over time, produces approximate-optimal solutions to games of ordered preference in receding horizon fast enough to potentially be practical. Lexicographic IBR over time equips autonomous agents in complex traffic scenarios with the ability to adapt to the environment in real time while ranking their preferences and considering those of others.

VI. ACKNOWLEDGMENTS

P. de las Heras Molins and E. Roy-Almonacid are partially supported by the E.U. Digital Europe Programme under grant DIGITAL-101083531-MERIT.

D. H. Lee and D. Fridovich-Keil are partially supported by the U.S. National Science Foundation (NSF) under Grant No. 2336840.

G. Bakirtzis is partially supported by the academic and research chair *Architecture des Systèmes Complexes* through the following partners: Dassault Aviation, Naval Group, Dassault Systèmes, KNDS France, Agence de l'Innovation de Défense, and Institut Polytechnique de Paris.

REFERENCES

- [1] D. M. Hausman, "Mistakes about preferences in the social sciences," *Philosophy of the Social Sciences*, 2011. DOI: 10.1177/0048393110387885.
- [2] A. Zanardi, G. Zardini, S. Srinivasan, S. Bolognani, A. Censi, F. Dörfler, and E. Frazzoli, "Posetal games: Efficiency, existence, and refinement of equilibria in games with prioritized metrics," *IEEE Robotics and Automation Letters*, 2022. DOI: 10.1109/LRA.2021.3135030.
- [3] J. M. M. Filho, E. Lucet, and D. Filliat, "Real-time distributed receding horizon motion planning and control for mobile multi-robot dynamic systems," in *ICRA*, 2017. DOI: 10.1109/ICRA.2017.7989081.
- [4] D. H. Lee, L. Peters, and D. Fridovich-Keil, "You can't always get what you want: Games of ordered preference," *IEEE Robotics and Automation Letters*, 2025. DOI: 10.1109/LRA.2025.3575324.
- [5] S. Hall, G. Belgioioso, D. Liao-McPherson, and F. Dörfler, "Receding horizon games with coupling constraints for demand-side management," in *CDC*, 2022. DOI: 10.1109/CDC51059.2022.9992497.
- [6] S. Hall, D. Liao-McPherson, G. Belgioioso, and F. Dörfler, *Stability certificates for receding horizon games*, arXiv:2404.12165 [eess.SY], 2024. DOI: 10.48550/ARXIV.2404.12165.
- [7] A. Censi, K. Slutsky, T. Wongpiromsarn, D. S. Yershov, S. Pendleton, J. G. M. Fu, and E. Frazzoli, "Liability, ethics, and culture-aware behavior specification using rulebooks," in *ICRA*, 2019. DOI: 10.1109/ICRA.2019.8794364.
- [8] T. Wongpiromsarn, K. Slutsky, E. Frazzoli, and U. Topcu, "Minimum-violation planning for autonomous systems: Theoretical and practical considerations," in *ACC*, 2021. DOI: 10.23919/ACC50511.2021.9483174.
- [9] M. Penlington, A. Zanardi, and E. Frazzoli, *Optimization of rulebooks via asymptotically representing lexicographic hierarchies for autonomous vehicles*, arXiv:2409.11199 [cs.RO], 2024. DOI: 10.48550/ARXIV.2409.11199.
- [10] A. Nurkanović, A. Pozharskiy, and M. Diehl, "Solving mathematical programs with complementarity constraints arising in nonsmooth optimal control," *Vietnam Journal of Mathematics*, 2024. DOI: 10.1007/s10013-024-00704-z.
- [11] Z. Qin, A. Ji, Z. Sun, G. Wu, P. Hao, and X. Liao, "Game theoretic application to intersection management: A literature review," *IEEE Transactions on Intelligent Vehicles*, 2024. DOI: 10.1109/TIV.2024.3379986.
- [12] F. Laine, D. Fridovich-Keil, C. Chiu, and C. J. Tomlin, "Multi-hypothesis interactions in game-theoretic motion planning," in *ICRA*, 2021. DOI: 10.1109/ICRA48506.2021.9561695.
- [13] Z. Williams, J. Chen, and N. Mehr, "Distributed potential iLQR: Scalable game-theoretic trajectory planning for multi-agent interactions," in *ICRA*, 2023. DOI: 10.1109/ICRA48891.2023.10161176.
- [14] N. Mehr, M. Wang, M. Bhatt, and M. Schwager, "Maximum-entropy multi-agent dynamic games: Forward and inverse solutions," *IEEE Transactions on Robotics and Automation*, 2023. DOI: 10.1109/TRO.2022.3232300.
- [15] E. L. Zhu and F. Borrelli, "A sequential quadratic programming approach to the solution of open-loop generalized Nash equilibria," in *ICRA*, 2023. DOI: 10.1109/ICRA48891.2023.10160799.
- [16] H. B. Jond, "Control of vehicle platoons with collision avoidance using noncooperative differential games," in *ITSC*, 2022. DOI: 10.1109/ITSC57777.2023.10422304.
- [17] L. Peters, D. Fridovich-Keil, L. Ferranti, C. Stachniss, J. Alonso-Mora, and F. Laine, "Learning mixed strategies in trajectory games," in *RSS*, 2022. DOI: 10.15607/RSS.2022.XVIII.051.
- [18] X. Liu, L. Peters, and J. Alonso-Mora, "Learning to play trajectory games against opponents with unknown objectives," *IEEE Robotics and Automation Letters*, 2023. DOI: 10.1109/LRA.2023.3280809.
- [19] K. Li, Y. Chen, M. Shan, J. Li, S. Worrall, and E. M. Nebot, "Game theory-based simultaneous prediction and planning for autonomous vehicle navigation in crowded environments," in *ITSC*, 2022. DOI: 10.1109/ITSC57777.2023.10422696.
- [20] L. Zhang, S. Han, and S. Grammatico, "An efficient game-theoretic planner for automated lane merging with multi-modal behavior understanding," in *ITSC*, 2022. DOI: 10.1109/ITSC57777.2023.10422316.
- [21] S. Veer, K. Leung, R. K. Cosner, Y. Chen, P. Karkus, and M. Pavone, "Receding horizon planning with rule hierarchies for autonomous vehicles," in *ICRA*, 2023. DOI: 10.1109/ICRA48891.2023.10160622.
- [22] J. F. Fisac, E. Bronstein, E. Stefansson, D. Sadigh, S. S. Sastry, and A. D. Dragan, "Hierarchical game-theoretic planning for autonomous vehicles," in *ICRA*, 2019. DOI: 10.1109/ICRA.2019.8794007.
- [23] F. J. Christophersen, "Receding horizon control," in *Optimal Control of Constrained Piecewise Affine Systems*. 2007. DOI: 10.1007/978-3-540-72701-9_3.
- [24] [SW] L. Peters, *ParametricMCPs.jl* 2025, URL: <https://github.com/JuliaGameTheoreticPlanning/ParametricMCPs.jl>.
- [25] D. Fridovich-Keil, *Smooth Game Theory*. 2024. [Online]. Available: https://clearoboticslab.github.io/documents/smooth_game_theory.pdf.
- [26] M. J. Kochenderfer and T. A. Wheeler, *Algorithms for Optimization*. The MIT Press, 2019.
- [27] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Review*, 2017. DOI: 10.1137/141000671.
- [28] [SW] L. Peters and A.-C. Papuc, *TrajectoryGamesBase.jl* 2025, URL: <https://github.com/JuliaGameTheoreticPlanning/TrajectoryGamesBase.jl>.
- [29] S. Dirkse and M. Ferris, "The path solver: A nonmonotone stabilization scheme for mixed complementarity problems," *Optimization Methods and Software*, 1995. DOI: 10.1080/10556789508805606.
- [30] [SW] P. de las Heras Molins, E. Roy-Almonacid, D. H. Lee, L. Peters, D. Fridovich-Keil, and G. Bakirtzis, *LEXIBROverTime.jl* 2025, DOI: 10.5281/zenodo.15825235.