Delft University of Technology

Delft Center for Systems and Control

Technical report 10-061

Stackelberg equilibria for discrete-time dynamic games – Part I: Deterministic games*

K. Staňková and B. De Schutter

If you want to cite this report, please use the following reference instead:
K. Staňková and B. De Schutter, “Stackelberg equilibria for discrete-time dynamic games – Part I: Deterministic games,” *Proceedings of the 2011 IEEE International Conference on Networking, Sensing and Control*, Delft, The Netherlands, pp. 249–254, Apr. 2011.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

Stackelberg Equilibria for Discrete-Time Dynamic Games Part I: Deterministic Games

Kateřina Staňková, Bart De Schutter

***Abstract*— We consider a two-person discrete-time dynamic game with the prespecified fixed duration. Each player maximizes her profit over the game horizon, taking decisions of the other player into account. Our goal is to find the Stackelberg equilibria for such a game. The solution approach differs with respect to the information available to individual players. While in the game with open-loop information structure the solution procedure is straightforward and already reported in the literature, the problem with the closed-loop problem information structure is difficult to solve, especially if twice differentiability of the leader's strategy is not imposed a priori. In this paper we focus on deterministic problems. We review classical optimization methods that can be used to solve the games with open-loop information structure. Additionally, we propose new methods for solving the games with the closed-loop information structure. Application of such methods is shown on specific examples. In the companion paper (Stackelberg Equilibria for Discrete-Time Dynamic Games – Part II: Stochastic Games with Deterministic Information Structure) we will consider a stochastic variant of the problem.**

Keywords: discrete-time infinite dynamic games, Stackelberg games, information structure, team problems

I. Introduction & literature overview

This paper deals with a deterministic variant of a two-person discrete-time infinite dynamic game with a prespecified duration. The game is referred to as infinite, because the decision spaces of the players comprise an infinite number of alternatives. We focus on the noncooperative variant of this game [1], [2], in which the goals of individual players might be conflicting. More specifically, we deal with Stackelberg problems [1]–[6], in which there exists hierarchy between individual players, as opposed to Nash problems [7], [8].

The open-loop Stackelberg solution concept in the infinite discrete-time dynamic games was first treated in [9]. Some other references that discuss the open-loop and the feedback Stackelberg solutions in discrete-time infinite dynamic games are [3], [10], [11]. Applications of this concept in microeconomics can be found in [12]. In this paper we review existing approaches applicable to solve the problem and relate them to other concepts, like dynamic programming [13] or the maximum principle [14].

Derivation of the global closed-loop Stackelberg solution of infinite discrete-time dynamic games remains a challenge. The global closed-loop Stackelberg solution for a specific classes of games with linear state dynamics and quadratic cost functional were found in [15], [16], while it was assumed a priori that such a solution is continuously differentiable. In this paper we concentrate on problems whose solutions may be non-differentiable or even discontinuous and we propose methods solving such problems. More specifically, we focus an indirect approach involving the team maximum of the game, similar to the approach proposed for linear-quadratic problems in [2], [17].

Due to space limitations we will not focus on another interesting problem: finding the solution of the game under feedback information structure [11]. Similarly the extension into the direction of the so-called inverse Stackelberg games [6], [18]–[20] is omitted in this paper, while this extension is discussed in the companion paper.

This paper is composed as follows. In Section II basic notions are introduced. In Section III the open-loop variant of the game is dealt with. In Section IV we study the closed-loop variant of the game. In Section V the conclusions and possibilities for future research are discussed.

II. Preliminaries

A. Basics

Let us consider a two-player Stackelberg game, in which P_1 is the leader and P_2 is the follower. Let $\mathscr{K} = \{1,\dots,K\}$, $K \in \mathbb{N}$ denote the stages of the game. Let $X \stackrel{\text{def}}{=} X^{(1)} \times X^{(2)}$ be the state space of the game with nonempty closed real intervals $X^{(1)}$, $X^{(2)}$. The state of the game for the k-th stage is then referred to as $x_k = \left(x_k^{(1)}, x_k^{(2)}\right)^T$, with P_1's state $x_k^{(1)}$ and P_2's state $x_k^{(2)}$. Let $U_k^{(1)} \subset \mathbb{R}$ (for each $k \in \mathscr{K}$) be a closed interval called the P_1's decision space. Its elements are P_1's permissible decisions $u_k^{(1)}$ at stage k, announced by P_1 at the beginning of each stage. Let $U_k^{(2)} \subset \mathbb{R}$ ($k \in \mathscr{K}$) be the P_2's decision space. Its elements are P_2's permissible decisions $u_k^{(2)}$ at stage k. The state of the game evolves according to the equation

$$x_{k+1} = f_k\left(x_k, u_k^{(1)}, u_k^{(2)}\right), \quad k \in \mathscr{K}, \tag{1}$$

with the initial state $x_1 \in X$ and the state function $f_k : X \times U_k^{(1)} \times U_k^{(2)} \to X$. Let the information gained and recalled by P_i at stage k of the game be determined by $\eta_k^{(i)}$, an a priori known selection from $\left(x_1^{(1)},\dots,x_k^{(1)};\, x_1^{(2)},\dots,\, x_k^{(2)}\right)$. Specifications of $\eta_k^{(i)}$ for all $k \in \mathscr{K}$ characterize the information structure of the game for P_i. Let $N_k^{(i)} \stackrel{\text{def}}{=} \left\{\left(\eta_k^{(i)}\right)^T\right\}$. Let $\Gamma_k^{(i)}$ be a prespecified class of measurable mappings $\gamma_k^{(i)} : N_k^{(i)} \to U_k^{(i)}$, called P_i's permissible strategies at stage k. The aggregate mapping $\gamma^{(i)} = \left(\gamma_1^{(i)}, \gamma_2^{(i)}, \dots, \gamma_K^{(i)}\right)$ is P_i's

K. Staňková and B. De Schutter are with the Delft Center for Systems & Control, Delft University of Technology, The Netherlands. `k.stankova@tudelft.nl`, `b.deschutter@tudelft.nl`

strategy, and the class $\Gamma^{(i)}$ of all such mappings $\gamma^{(i)}$ so that $\gamma_k^{(i)} \in \Gamma_k^{(i)}$, $k \in \mathscr{K}$, is P_i's strategy set. We will refer to a P_2's strategy based on the P_1's strategy as the P_2's response to the P_1's strategy.

Definition 2.1: (Information structure) *In a two-person discrete time dynamic game, we say that P_i's ($i \in \{1,2\}$) information has (for all $k \in \mathscr{K}$)*

(a) *an open-loop structure if* $\eta_k^{(i)} = x_1$,
(b) *a closed-loop structure if* $\eta_k^{(i)} = (x_1, \ldots, x_k)$,
(c) *a feedback structure if* $\eta_k^{(i)} = x_k$.

In this paper we will consider cases (a) and (b), while case (c) is a subject of the future research.

Let $L^{(i)} : \left(X \times U_1^{(1)} \times U_1^{(2)}\right) \times \left(X \times U_2^{(1)} \times U_2^{(2)}\right) \times \ldots \times \left(X \times U_k^{(1)} \times U_k^{(2)}\right) \to \mathbb{R}$ be called P_i's profit function. Each player maximizes $L^{(i)}$, taking into account possible actions of the other player.

B. Assumptions on state and profit functions

In order to simplify the analysis, we will mostly assume (unless stated differently), the following: The functions $f_k\left(\cdot, u_k^{(1)}, u_k^{(2)}\right)$, $f_k\left(x_k, \cdot, u_k^{(2)}\right)$, and $f_k\left(\cdot, u_k^{(1)}, \cdot\right)$ are continuously differentiable on $\mathbb{R}_+^2$, $U_k^{(1)}$, and $U_k^{(2)}$, respectively. The profit functions are stage-additive, i.e., there exists $g_k^{(i)} : X^2 \times U_k^{(1)} \times U_k^{(2)} \to \mathbb{R}$, for all $k \in \mathscr{K}$, so that[1]

$$L^{(i)}\left(u^{(1)}, u^{(2)}\right) = \sum_{k=1}^{K} g_k^{(i)}\left(x_k, u_k^{(1)}, u_k^{(2)}, x_{k+1}\right), \qquad (2)$$

$i \in \{1,2\}$, and continuously differentiable on $U^{(1)} \times U^{(2)}$. Furthermore, $g_k^{(i)}\left(\cdot, u_k^{(1)}, u_k^{(2)}, \cdot\right)$ and $g_k^{(i)}\left(x_k, u_k^{(1)}, \cdot, x_{k+1}\right)$ are continuously differentiable on $\mathbb{R}^2$ and $U_k^{(2)}$, respectively.

These assumptions might be too restrictive for real-world applications, but they assure us the existence and with some additional assumptions also uniqueness of the solutions of the problems dealt with in this paper. However, we will discuss how to proceed in more general situations when applicable.

C. Game formulation

An extensive form description of the game contains the set of players, the index set defining the stages of the game, the state space and the decision spaces, the state equation, the observation sets, the state-observation equation, the information structure of the game, the information spaces, the strategy sets, and the profit functionals.

Similarly as it is done in [2] for a general N-person discrete-time game, we can transform the game into a normal-form game. For each fixed initial state x_1 and for each pair $\left(\gamma^{(1)}, \gamma^{(2)}\right)$, where $\gamma^{(i)} \in \Gamma^{(i)}$, $i \in \{1,2\}$, the extensive form description leads to the unique vector sequence

$$\left(u_k^{(i)} \stackrel{\text{def}}{=} \gamma_k^{(i)}\left(\eta_k^{(i)}\right), \quad i \in N, \quad k \in \mathscr{K}\right). \qquad (3)$$

Then, substitution of (3) into $L^{(1)}$ and $L^{(2)}$ leads to a pair of functions reflecting the corresponding profits of the players.

[1]For the sake of notation convenience we refer to the profit functions as $L^{(i)}\left(u^{(1)}, u^{(2)}\right)$, with $u^{(i)} \stackrel{\text{def}}{=} \left(u_1^{(i)}, \ldots, u_K^{(i)}\right)^T$.

This further implies existence of a composite mapping $J^{(i)} : \Gamma^{(1)} \times \Gamma^{(2)} \to \mathbb{R}$ for each $i \in \{1,2\}$, which is the strategy-dependent profit function. The permissible strategy spaces $\left(\Gamma^{(1)}, \Gamma^{(2)}\right)$ and the pair $\left(J^{(1)}(\gamma^{(1)}, \gamma^{(2)}), J^{(2)}(\gamma^{(1)}, \gamma^{(2)})\right)$ constitute the normal form description of the game for each fixed initial state x_1. Under the normal form description there is no essential difference between infinite discrete-time dynamic games and finite games. This means that techniques used for finding solutions of finite games, such as saddle-point, Nash, and Stackelberg equilibrium solution concepts, introduced originally for finite games, can be used for solving infinite discrete-time dynamic games [2].

III. Open-loop game

In this section we summarize and generalize already known results on Stackelberg equilibria for games with the open-loop information structure, which can be found in [2], [3], [10], [11], [21]. To solve such problems, the classical techniques of optimal control theory, i.e., the maximum principle [14], [22] and dynamic programming [2], [13], [23], can be used. By recursive substitution of (1) into (2), the profit functions can be made dependent only on $u^{(1)}$ and $u^{(2)}$, and x_1, which is known a priori. Then the game can be viewed as a static game.

Definition 3.1: (P_2's optimal response) The set $R^{(2)}\left(\gamma^{(1)}\right) \subset \Gamma^{(2)}$ defined for each $\gamma^{(1)} \in \Gamma^{(1)}$ by $R^{(2)}\left(\gamma^{(1)}\right) = \left\{\zeta \in \Gamma^{(2)} | J^{(2)}\left(\gamma^{(1)}, \zeta\right) \geq J^{(2)}\left(\gamma^{(1)}, \gamma^{(2)}\right), \ \forall \gamma^{(2)} \in \Gamma^{(2)}\right\}$ is P_2's optimal response to a strategy $\gamma^{(1)} \in \Gamma^{(1)}$ of P_1.

Definition 3.2: (Stackelberg equilibrium strategy) In a Stackelberg game with P_1 as the leader and P_2 as the follower, a strategy $\gamma^{(1)*}$ is P_1's Stackelberg equilibrium strategy, if $\min_{\gamma^{(2)} \in R^{(2)}(\gamma^{(1)*})} J^{(1)}\left(\gamma^{(1)*}, \gamma^{(2)}\right)$
$= \max_{\gamma^{(1)} \in \Gamma^{(1)}} \min_{\gamma^{(2)} \in R^{(2)}(\gamma^{(1)})} J^{(1)}\left(\gamma^{(1)}, \gamma^{(2)}\right) \stackrel{\text{def}}{=} J^{(1)*}$. The quantity $J^{(1)*}$ is P_1's profit when she plays the Stackelberg strategy.

Definition 3.3: (Stackelberg equilibrium solution) Let $\gamma^{(1)*} \in \Gamma^{(1)}$ be P_1's Stackelberg strategy. Then, any $\gamma^{(2)*} \in R^{(2)}\left(\gamma^{(1)*}\right)$ is P_2's optimal strategy that is in equilibrium with $\gamma^{(1)*}$. The pair $\left(\gamma^{(1)*}, \gamma^{(2)*}\right)$ is a Stackelberg game with P_1 as the leader, and $\left(J^{(1)}\left(\gamma^{(1)*}, \gamma^{(2)*}\right), \ J^{(2)}\left(\gamma^{(1)*}, \gamma^{(2)*}\right)\right)$ represents the corresponding Stackelberg outcomes.

When the conditions stated in Section II-B hold, the solution of the game always exists, but might be nonunique. The following proposition follows directly from Definition 3.2.

Proposition 3.1: If conditions stated in Section II-B hold, the Stackelberg equilibrium of the open-loop discrete dynamic game defined by (1) *and* (2) *is a singleton if* $L^{(2)}\left(u^{(1)}, \cdot\right)$ *is strictly concave on* $U^{(2)}$ *for all* $u_1 \in U^{(1)}$ and $L^{(1)}\left(\cdot, u^{(2)}\right)$ is strictly concave on $U^{(1)}$ for all $u_2 \in U^{(2)}$

Let us focus on the approaches which lead to finding a Stackelberg solution of the game defined by (1) and (2).

Approach 1

The standard way to find the optimal strategy for P_1 is to determine P_2's optimal response to P_1's decision by maximizing $L^{(2)}\left(u^{(1)}, u^{(2)}\right)$ for every fixed $u^{(1)} \in U^{(1)}$ [2]. Denoting this mapping by $\mathscr{D}: U^{(1)} \to U^{(2)}$, the optimization problem faced by P_1 is then maximization of $L^{(1)}\big(u^{(1)}, \mathscr{D}\big(u^{(1)}\big)\big)$ over $U^{(1)}$, yielding P_1's Stackelberg strategy in this open-loop game. With an increasing number of stages the dimension of vectors $u^{(1)}$, $u^{(2)}$ increases as well. Therefore, such a derivation of the P_1's strategies is not considered applicable if K is high.

Approach 2

Another option, more applicable for problems with high K, is to first determine P_2's optimal response to every strategy $\gamma^{(1)} \in \Gamma^{(1)}$ of P_1. If the functions f_k and g_k satisfy the conditions from Section II-B, there exists an optimal response $\overline{\gamma}^{(1)}(x_1) = \overline{u}^{(2)}$ of P_2 to any announced strategy $u^{(1)} = \gamma^{(1)} \in \Gamma^{(1)}$ of P_1, satisfying

$$\begin{aligned}
\overline{x}_{k+1} &= f_k\big(\overline{x}_k, u_k^{(1)}, \overline{u}_k^{(2)}\big), \quad \overline{x}_1 = x_1,\\
\overline{u}_k^{(2)} &= \arg\max_{u_k^{(2)} \in U_k^{(2)}} H_k^{(2)}(\lambda_{k+1}, u_k^{(1)}, u_k^{(2)}, \overline{x}_k),\\
\lambda_k &= \nabla_{x_k} f_k\big(\overline{x}_k, u_k^{(1)}, \overline{u}_k^{(2)}\big)^T\\
&\quad \cdot \big[\lambda_{k+1} + \big(\frac{\partial}{\partial x_{k+1}} g_k^{(2)}\big(\overline{x}_k, u_k^{(1)}, \overline{u}_k^{(2)}, \overline{x}_{k+1}\big)\big)^T\big]\\
&\quad + \big[\nabla_{x_k} g_k^{(2)}\big(\overline{x}_k, u_k^{(1)}, \overline{u}_k^{(2)}, \overline{x}_{k+1}\big)\big]^T; \quad \lambda_{K+1} = 0,\\
H_k^{(2)}\big(\lambda_{k+1}, u_k^{(1)}, u_k^{(2)}, x_k\big) &\overset{\text{def}}{=} g_k^{(2)}\big(x_k, u_k^{(1)}, u_k^{(2)}, f_k(x_k, u_k^{(1)}, u_k^{(2)})\big)\\
&\quad + \lambda_{k+1}^T f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big).
\end{aligned}$$

The sequence $\lambda_1, \ldots, \lambda_{K+1}$ is a sequence of two-dimensional costate vectors. To obtain P_1's optimal Stackelberg strategy, one has to maximize $L^{(1)}\big(u^{(1)}, u^{(2)}\big)$, taking into account P_2's optimal response. Player P_1 is then faced with the problem

$$\begin{aligned}
&\max_{u^{(1)} \in U^{(1)}} L^{(1)}\big(u^{(1)}, u^{(2)}\big), \quad \text{subject to}\\
&x_{k+1} = f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big), \quad x_1 \text{ given},\\
&\lambda_k = F_k\big(x_k, u_k^{(1)}, u_k^{(2)}, \lambda_{k+1}\big), \quad \lambda_{K+1} = 0,\\
&\frac{\partial}{\partial u_k^{(2)}} H_k^{(2)}\big(\lambda_{k+1}, u_k^{(1)}, u_k^{(2)}, x_k\big) = 0 \quad (k \in \mathscr{K}),\\
&\text{where} \quad F_k \overset{\text{def}}{=} \nabla_{x_k} f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big)^T \lambda_{k+1} +\\
&+ \Big[\nabla_{x_k} g_k^{(2)}\big(x_k, u_k^{(1)}, u_k^{(2)}, f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big)\big)\Big]^T.
\end{aligned}$$

Proposition 3.2: *If $\big(\gamma_k^{(1)*}\big(x_1\big) = u_k^{(1)*} \in U_k^{(1)}, k \in \mathscr{K}\big)$ denotes an open-loop Stackelberg equilibrium in the dynamic game, there exist finite vector sequences $\lambda_1, \ldots, \lambda_K$, $\mu_1, \ldots, \mu_K$, $\nu_1, \ldots, \nu_K$, that satisfy the following relations:*

$$\begin{aligned}
x_{k+1}^* &= f_k\big(x_k^*, u_k^{(1)*}, u_k^{(2)*}\big), \quad x_1^* = x_1,\\
&\frac{\partial}{\partial u_k^{(1)}} H_k^{(1)}\big(\lambda_k, \mu_k, \nu_k, p_{k+1}^*, u_k^{(1)*}, u_k^{(2)*}, x_k^*\big) = 0,\\
&\frac{\partial}{\partial u_k^{(2)}} H_k^{(1)}\big(\lambda_k, \mu_k, \nu_k, p_{k+1}^*, u_k^{(1)*}, u_k^{(2)*}, x_k^*\big) = 0,\\
\lambda_{k-1}^T &= \nabla_{x_k} H_k^{(1)}\big(\lambda_k, \mu_k, \nu_k, p_{k+1}^*, u_k^{(1)*}, u_k^{(2)*}, x_k^*\big), \quad \lambda_K = 0,\\
\mu_{k+1}^T &= \frac{\partial}{\partial \lambda_{k+1}} H_k^{(1)}\big(\lambda_k, \mu_k, \nu_k, p_{k+1}^*, u_k^{(1)*}, u_k^{(2)*}, x_k^*\big), \quad \mu_1 = 0,\\
&\frac{\partial}{\partial u_k^{(2)}} H_k^{(2)}\big(\lambda_{k+1}^*, u_k^{(1)*}, u_k^{(2)*}, x_k^*\big) = 0.\\
\lambda_k^* &= F_k\big(x_k^*, u_k^{(1)*}, u_k^{(2)*}, \lambda_{k+1}^*\big), \quad \lambda_{K+1}^* = 0,\\
\textit{where} \quad H_k^{(1)} &= g_k^{(1)}\big(x_k, u_k^{(1)}, u_k^{(2)}, f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big)\big)\\
&\quad + \mu_k^T F_k\big(x_k, u_k^{(1)}, u_k^{(2)}, \lambda_{k+1}\big) + \lambda_k^T f_k\big(x_k, u_k^{(1)}, u_k^{(2)}\big)\\
&\quad + \nu_k^T \big(\frac{\partial}{\partial u_k^{(2)}} H_k^{(2)}\big(\lambda_{k+1}, u_k^{(1)}, u_k^{(2)}, x_k\big)\big)^T.
\end{aligned}$$

Furthermore, $\big(u_k^{(2)*} | k \in \mathscr{K}\big)$ is the corresponding unique open-loop Stackelberg strategy of P_2 and $\big(x_{k+1}^* | k \in \mathscr{K}\big)$ is the state trajectory associated with the Stackelberg solution.

Proof. The problem can be transformed into a finite nonlinear programming problem [11]. The proof then follows from application of the dynamic programming principle [2], [23], [24] to such a problem. □

Using Proposition 3.2, a closed-form solution of the open-loop Stackelberg problem can be found recursively.

Remark 3.1: (Closed-loop information structure for P_2) It can be shown that if P_2 has a closed-loop information, her optimal response will be any closed-loop representation of the open-loop policy $\big(\overline{u}_k^{(2)} | k \in \mathscr{K}\big)$, i.e., any strategy that will generate the same unique state trajectory and that will have the same open-loop value on this trajectory. However, P_1's unique optimal strategy will remain the same, whereas P_2's corresponding optimal response strategy may be nonunique.

Discussion

Proposition 3.1 discusses a very specific situation in which $L^{(2)}(u^{(1)}, \cdot)$ is strictly concave on $U^{(2)}$. In such a situation the optimal response of P_2 to any P_1's decision is unique. Even if P_1 has multiple sequences of decisions yielding the same profit for her, the decision among them yielding the lowest profit to P_2 is unique. Strict concavity of $L^{(1)}$ then yields the unique game solution.

Without assumptions defined in Section II-B the game does not need to have a classical solution, even if the Hamiltonians $H_k^{(1)}$ and $H_k^{(2)}$ are smooth. In such situations the problem can be approached by looking for generalized solutions, satisfying the conditions in Proposition 3.2 almost everywhere. Existence results for general Hamilton-Jacobi-Bellman equations have been obtained by several authors, e.g., [25]–[27], with the most general results being given by [28]. Carrying the steps equivalent to the steps introduced in this section to find a classical solution can then be carried out to find the class of generalized solutions. The question of uniqueness of the generalized solution is more complex. The notion of viscosity solution [29], [30], which may be nondifferentiable and for which uniqueness (and even stability and general existence) theorems are available, should be introduced if we wish to obtain the unique generalized solution.

IV. Closed-loop game

In this section we will first show the way how to find the feedback Stackelberg equilibrium and the problem of finding the (global) Stackelberg solution of the closed-loop Stackelberg game, the main problem dealt with in this paper.

A. Feedback Stackelberg equilibrium

With the closed-loop information structure the feedback solution of the game can be obtained recursively, using dynamic programming and solving a static Stackelberg game at each stage. Our aim is to obtain the feedback Stackelberg solution valid for all possible initial states $x_1 \in X$. A pair $\left(\gamma^{(1)*}, \gamma^{(2)*}\right)$ constitutes a feedback Stackelberg solution if for all appropriate x_k

$$\max_{\gamma_k^{(1)} \in \Gamma_k^{(1)}} \max_{\gamma_k^{(2)} \in R_k^{(2)}\left(\gamma_k^{(1)}\right)} \widehat{\Psi}_k^{(1)}\left(\gamma_k^{(1)}, \gamma_k^{(2)}, x_k\right) = \widehat{\Psi}_k^{(1)}\left(\gamma_k^{(1)*}, \gamma_k^{(2)*}, x_k\right),$$

where $R_k^{(2)}\left(\gamma_k^{(1)}\right)$ is a singleton set defined by $R_k^{(2)}\left(\gamma_k^{(1)}\right) = \left\{\chi \in \Gamma_k^{(2)} : \widehat{\Psi}_k^{(2)}\left(\gamma_k^{(1)}, \chi, x_k\right) = \max_{\gamma_k^{(2)} \in \Gamma_k^{(2)}} \widehat{\Psi}_k^{(2)}\left(\gamma_k^{(1)}, \gamma_k^{(2)}, x_k\right) \ \forall x_k \in \mathbb{R}_+ \times [0,1]\right\}$, $\widehat{\Psi}_k^{(i)}\left(\gamma_k^{(1)}, \gamma_k^{(2)}, x_k\right) \stackrel{\text{def}}{=} \Psi_k^{(i)}\left(f_k\left(x_k, \gamma_k^{(1)}(x_k), \gamma_k^{(2)}(x_k)\right), \gamma_k^{(1)}(x_k), \gamma_k^{(2)}(x_k), x_k\right)$, $i \in \{1,2\}$, $k \in \mathscr{K}$, and $\Psi_k^{(i)}$ can be defined recursively as $\Psi_k^{(i)}\left(x_{k+1}, \gamma_k^{(1)}(x_k), \gamma_k^{(2)}(x_k), x_k\right)$ $= \Psi_{k+1}^{(i)}\left(f_{k+1}\left(x_{k+1}, \gamma_{k+1}^{(1)*}(x_{k+1}), \gamma_{k+2}^{(2)*}(x_{k+1})\right), \gamma_{k+1}^{(1)*}(x_{k+1}), \gamma_{k+1}^{(2)*}(x_{k+1}), x_{k+1}\right) + g_k^{(i)}\left(x_k, \gamma_k^{(1)}(x_{k+1}), \gamma_k^{(2)}(x_k), x_{k+1}\right)$; $\Psi_{k+1}^{(i)} = 0$. With $\widehat{\Psi}_k^{(2)}\left(\gamma_k^{(1)}, \cdot, x_k\right)$ strictly concave on $U_k^{(2)}$ ($k \in \mathscr{K}$) and a singleton set $R_k^{(2)}$ the recursive definition of $\Psi_k^{(i)}$ provides an easily implementable procedure for computation of feedback Stackelberg strategies.

B. (Global) Stackelberg equilibrium

The (global) Stackelberg equilibrium in closed-loop decision problems cannot be found by utilizing standard optimal control techniques, as the reaction set of P_2 cannot be generally determined in a closed form for all possible strategies of P_1. To show the difficulties encountered when solving the problem in which P_1 has access to dynamic information and to motivate the solution approach that we propose, we will consider two case studies: In Section IV-B.1 we will deal with a 2-stage game with each P_1 and P_2 acting only once. If, additionally, P_2 acts in the last stage of the game, the problem becomes more difficult to solve. Therefore, in Section IV-B.2 such a problem will be dealt with. On the basis of these examples we propose the methodology to solve the general closed-loop games (Section IV-C).

1) First motivation problem: Let

$$\begin{aligned} &L^{(1)} = -3{x_3}^2 + \left(u_2^{(1)}\right)^2 - \alpha\left(u_1^{(2)}\right)^2, \quad 1 > \alpha \geq 0, \\ &L^{(2)} = -5{x_3}^2 - \left(u_1^{(2)}\right)^2, \\ &x_2 = x_1 + u_1^{(2)}, \quad x_3 = x_2 + u_2^{(1)}. \end{aligned} \tag{4}$$

We assume that P_1, acting at stage 2, has access to both x_1 and x_2, while P_2 has access to x_1 only.

To any strategy $\gamma^{(1)} \in \Gamma^{(1)}$ announced by P_1 an optimal reaction of P_2 equals

$$\arg\max_{u_1^{(2)} \in \mathbb{R}} \left(-5\left(x_1 + u_1^{(2)} + \gamma^{(1)}(x_2, x_1)\right)^2 - \left(u_1^{(2)}\right)^2\right).$$

Writing such a strategy symbolically as $\gamma^{(2)}\left(x_1; \gamma^{(1)}\right)$, the goal of P_1 is to find[2]

$$\begin{aligned} &\arg\max_{\gamma^{(1)} \in \Gamma^{(1)}} \left(-3\left(\gamma^{(2)}\left(x_1;\gamma^{(1)}\right) + \gamma^{(1)}\left(x_1 + \gamma^{(2)}\left(x_1;\gamma^{(1)}\right), x_1\right)\right.\right. \\ &\left. + x_1\right)^2 + \left(\gamma^{(1)}\left(x_1 + \gamma^{(2)}\left(x_1;\gamma^{(1)}\right), x_1\right)\right)^2 \\ &\left. - \alpha\left(\gamma^{(2)}\left(x_1;\gamma^{(1)}\right)\right)^2\right). \end{aligned} \tag{5}$$

This is a constrained optimization problem with the constraint in the form of the maximum of a function. If we restricted the permissible strategies of P_1 to C^2-functions of the first argument, we could obtain the first-order and the second-order conditions implicitly determining $\gamma^{(2)}\left(x_1, \gamma^{(1)}\right)$ [31]. However, depending on α, P_1's optimal strategy may be nondifferentiable. Therefore, we are interested in methods not relying on the twice differentiability of P_1's optimal strategy.

The problem can be approached by looking for generalized solutions, satisfying (4) and (5) almost everywhere [28]. However, the class of generalized solutions can be large and choosing the appropriate elements of this class may require nontrivial analysis. The notion of viscosity solution [29], [30] should be introduced if we wish to obtain the unique generalized solution. However, this approach may cause difficulties if there exist multiple classical solutions of the problem [30]. We need to adopt an approach which leads to finding both classical and generalized solutions.

Therefore, in this paper we propose another approach, involving the so-called *team maximum*

$$\max_{\gamma^{(1)} \in \Gamma^{(1)}} \max_{\gamma^{(2)} \in \Gamma^{(2)}} J^{(1)}\left(\gamma^{(1)}, \gamma^{(2)}\right). \tag{6}$$

This value is clearly an upper bound on the profit of P_1. Finding (6) is referred to as the *team problem* [2], [6]. To find strategies that imply (6) we utilize a dynamic programming approach [13], [23]. The pair of feedback strategies

$$\gamma^{(1),\dagger}(x_2) = -\frac{3x_2}{2}, \quad \gamma^{(2),\dagger}(x_1) = -\frac{3x_1}{3-2\alpha} \tag{7}$$

provides the unique globally maximizing solution within the class of feedback strategies. The corresponding optimal state trajectory is described by

$$x_2^\dagger = -\frac{2x_1\alpha}{3-2\alpha}, \quad x_3^\dagger = -\frac{x_2^\dagger}{2}. \tag{8}$$

and the team maximum for P_1 is

$$J^{(1)}\left(\gamma^{(1),\dagger}, \gamma^{(2),\dagger}\right) = -\frac{3\alpha {x_1}^2}{3-2\alpha}. \tag{9}$$

If a game extension, in which P_1 knows x_1, is considered, $\gamma^{(1)}$ becomes a nonunique optimal strategy for P_1, but any representation of this strategy on the state trajectory (8) also constitutes an optimal strategy. More generally, if we denote the class of such strategies by $\Gamma^{(1),\dagger}$, any pair

$$\gamma^{(1),\dagger} \in \Gamma^{(1),\dagger}, \quad \gamma^{(2),\dagger}(x_1) = -\frac{3x_1}{3-2\alpha} \tag{10}$$

[2]If such a solution is nonunique, we take a supremum of all such solutions.

constitutes a team-maximum solution. For each such pair, the state trajectory and the corresponding profit are still given by (8) and (9), respectively.

For (6) to be realized, there should exist an element of $\Gamma^{(1),\dagger}$, which we will denote by $\gamma^{(1)*}$, which forces P_2 to choose $\gamma^{(2),\dagger}$ even though she is maximizing her own profit, i.e., $\gamma^{(2),\dagger} \stackrel{\text{def}}{=} \arg\max_{\gamma^{(2)}\in\Gamma^{(2)}} J^{(2)}\left(\gamma^{(1)*},\gamma^{(2)}\right)$. Moreover, the maximum of $J^{(2)}\left(\gamma^{(1),\dagger},\cdot\right)$ on $\Gamma^{(2)}$ has to be obtained uniquely at $\gamma^{(2),\dagger}$.[1] Intuitively, an optimal strategy for P_1 is the one that implies $\gamma^{(1),\dagger}$ from (7) if P_2 plays $\gamma^{(2),\dagger}$ from (7) and that penalizes P_2 otherwise. One of such strategies is

$$\gamma^{(1),\dagger}(x_1,x_2) = \begin{cases} \rho x_1, & \text{if } \frac{x_2}{x_1} \neq -\frac{2\alpha}{3-2\alpha}, \\ -\frac{3x_2}{2}, & \text{if } \frac{x_2}{x_1} = -\frac{2\alpha}{3-2\alpha}. \end{cases} \tag{11}$$

Therefore, (11) constitutes a (discontinuous) Stackelberg strategy for P_1, with the unique optimal response of P_2 being $\gamma^{(2),\dagger}$ from (7). It can be shown that for this game under the closed-loop information structure a Stackelberg solution exists for all $\alpha \geq 0$. The Stackelberg strategy of P_1 is nonunique, but the optimal response of P_2 is unique.

Lemma 4.1: The Stackelberg solutions of the game from Example IV-B.1 constitute Nash equilibrium solutions.

Proof. A Stackelberg solution satisfies $J^{(2)}\left(\gamma^{(1),\dagger},\gamma^{(2),\dagger}\right) \geq J^{(2)}\left(\gamma^{(1),\dagger},\gamma^{(2)}\right), \quad \forall\gamma^{(2)} \in \Gamma^{(2)}$. The inequality $J^{(1)}\left(\gamma^{(1),\dagger},\gamma^{(2),\dagger}\right) \geq J^{(2)}\left(\gamma^{(1)},\gamma^{(2),\dagger}\right), \quad \forall\gamma^{(1)} \in \Gamma^{(1)}$ also holds, since the Stackelberg solution is also team-optimal in this case. □

We can recapitulate the outcomes of the example:

- The Stackelberg profit of P_1 is equal to her team maximum.
- A Stackelberg solution exists for all $\alpha > 0$.
- The Stackelberg strategy of P_1 is nonunique (depends, in this example, on ρ), but the optimal response of P_2 to all those strategies is nonunique (and independent of ρ).
- Optimal P_1's strategy may be discontinuous.

Team-maximum based approach led to finding a closed-loop solution. This approach can be recursively applied to games with more stages and more complex dynamics, but the related team problem is not always the one determined by (6), especially if P_2 acts in the last stage of the game. In the following example we show how to deal with such situations.

2) Second motivation problem: Let

$$\begin{aligned} L^{(1)} &= -3{x_3}^2 + \left(u_2^{(1)}\right)^2 - \alpha\left(u_1^{(2)}\right)^2, \quad 1 > \alpha \geq 0, \\ L^{(2)} &= -5{x_3}^2 - \left(u_1^{(2)}\right)^2 - \left(u_2^{(2)}\right)^2, \\ x_2 &= x_1 + u_1^{(2)}, \quad x_3 = x_2 + u_2^{(1)} + u_2^{(2)}. \end{aligned} \tag{12}$$

We assume that P_1 acting at stage 2 has a single decision variable $u_1^{(2)}$, while P_2 acting in both stages 1 and 3 has decision variables $u_1^{(2)}$ and $u_2^{(2)}$. The optimal reaction of P_2 to any announced strategy $\gamma^{(1)}$ of P_1 is $\gamma_2^{(2)}\left(x_2;\gamma^{(1)}\right) = -\frac{5}{6}\left(x_2 + \gamma^{(1)}(x_2,x_1)\right)$. Substituting this expression into $J^{(1)}$ and $J^{(2)}$ derived from $L^{(1)}$ and $L^{(2)}$, respectively, gives

$$\begin{aligned} \tilde{J}^{(1)} &= -\frac{1}{12}{x_2}^2 - \frac{1}{6}x_2\gamma^{(1)}(x_2,x_1) + \frac{11}{12}\left(\gamma^{(1)}(x_2,x_1)\right)^2 \\ &\quad - \alpha\left(\gamma^{(2)}(x_1)\right)^2, \\ \tilde{J}^{(2)} &= -\frac{5}{6}{x_2}^2 - \frac{5}{3}x_2\gamma^{(1)}(x_2,x_1) \\ &\quad - \frac{5}{6}\left(\gamma^{(1)}(x_2,x_1)\right)^2 - \left(\gamma^{(2)}(x_1)\right)^2. \end{aligned}$$

The maximum that P_1 can achieve is then

$$\max_{\gamma^{(1)}\in\Gamma^{(1)}} \max_{\gamma^{(2)}\in\Gamma^{(2)}} \tilde{J}^{(1)}\left(\gamma^{(1)},\gamma(2)\right). \tag{13}$$

By this way we have converted the problem into a different problem in which P_1 does not act in the last stage, similar to the example in Section IV-B.1, with the corresponding team problem (13). The problem solution in feedback strategies is

$$\gamma^{(1),\dagger} = \frac{x_2}{11}, \quad \gamma^{(2),\dagger} = -\frac{x_1}{11\alpha+1}, \tag{14}$$

with the corresponding unique trajectory being

$$x_2^\dagger = \frac{11x_1\alpha}{11\alpha+1}, \quad x_3^\dagger = \frac{2x_2^\dagger}{11}. \tag{15}$$

The associated team maximum is $-\frac{\alpha {x_1}^2}{1+11\alpha}$, which is an upper bound on the P_1's profit and which is lower than the team maximum of the original game (12). We proceed as in Example IV-B.1 and obtain a strategy for P_1 that forces P_2 to strategy $\gamma^{(2),\dagger}$ from (14):

$$\gamma^{(1),\dagger}(x_1,x_2) = \begin{cases} \rho x_1, & \text{if } \frac{x_2}{x_1} \neq \frac{11\alpha}{11\alpha+1}, \\ \frac{x_2}{11}, & \text{if } \frac{x_2}{x_1} = \frac{11\alpha}{11\alpha+1}, \end{cases} \tag{16}$$

with $\rho > 0$. The optimal response of P_2 to such a strategy is $\gamma_1^{(2)*}(x_1) = -\frac{x_1}{11\alpha+1}$, $\gamma_2^{(2)*}(x_1,x_2) = -\frac{5}{6}\left(x_2 + \gamma^{(1)*}(x_2,x_1)\right)$, bringing to P_1 profit (13).

It can be shown that for $\alpha = 0$ no continuously differentiable Stackelberg equilibrium exists. Moreover, because P_2 acts at the last stage of the game, there does not exist a Stackelberg solution that would be a Nash equilibrium solution.

C. General approach

Let us now consider a game with closed-loop information structure, dynamics defined by (1), and profit functions (2). Based on the analysis in Sections IV-B.1 and IV-B.2, we propose the following approach:

- If the follower acts in the last stage of the game K, convert the game into the game with the leader acting last, as was shown in Section IV-B.2.
- Find the team maximum of the game for the leader, which determines the upper bound of the leader's profit.
- Adopt a particular representation of the optimal team strategy of the leader in this team problem. This particular strategy must force the follower to maximize the profit function of the team problem while she is maximizing her own profit function. Such a strategy can be found following the example in Section IV-B.1.

As was it shown in Sections IV-B.1 and IV-B.2, the leader's strategy may be nonunique and may be discontinuous. However, if there exists a smooth optimal strategy of P_1, then this strategy can be also found using the indirect approach presented in this section. For some problems a P_1's strategy forcing the follower to maximize the P_1's profit function cannot be found. This does not need to mean that there is no team-optimal solution. However, if we are unable to find such a solution, we can restrict ourselves to suboptimal solutions of the original game, as it will be shown in the companion paper for the stochastic variant of the game (1)-(2).

V. Conclusions & Future research

We have introduced specific types of discrete-time infinite dynamic games and have proposed methods to find their Stackelberg equilibrium solutions. Such solutions depend on the information structure of the games. We have reviewed already existing methods applicable to solve problems with open-loop information structure. After having studied two specific examples we have proposed an indirect method to solve the game with closed-loop information structure for general problems. This method is applicable especially if the classical solutions of the problem may not exist.

While in this paper we have focused on deterministic problems, in the companion paper we will focus on their stochastic variants.

Future research consists of extending the approaches proposed here to feedback Stackelberg games and inverse Stackelberg problems.

Acknowledgments

This research has been supported by the European 7th framework STREP project "Hierarchical and distributed model predictive control of large-scale systems (HD-MPC)", contract number INFSO-ICT-223854, and by the European 7th Framework Network of Excellence "Highly-complex and networked control systems (HYCON2)". The authors would like to thank dr. Ján Buša from the Department of Mathematics, Faculty of Electrical Engineering & Computer Science, The Technical University of Košice, for his helpful insights.

References

[1] A. Perea y Monsuwé, *Information Structures in Non-Cooperative Games*. Maastricht, The Netherlands: Unigraphic, 1997.

[2] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, Pennsylvania: SIAM, 1999.

[3] F. Kydland, "Noncooperative and dominant player solutions in discrete dynamic games," *International Journal of Economic Review*, vol. 16, pp. 321–335, 1975.

[4] A. Bagchi, *Stackelberg Differential Games in Economical Models*. Berlin, Germany: Springer-Verlag, 1984.

[5] H. Peters, *Game Theory: A Multi-Leveled Approach*. Dordrecht, The Netherlands: Springer, 2008.

[6] K. Staňková, "On Stackelberg and Inverse Stackelberg Games & Their Applications in the Optimal Toll Design Problem, the Energy Market Liberalization Problem, and in the Theory of Incentives," Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 2009.

[7] J. Nash, "Noncooperative games," *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.

[8] G. Olsder, "Adaptive Nash strategies for repeated games resulting in Pareto solutions," Delft University of Technology, Department of Mathematics and Informatics, Reports of the Department of Mathematics and Informatics 86-09, 1986.

[9] M. Simaan and J. B. Cruz, "Additional aspects of the Stackelberg strategy in the nonzero sum games," *Journal of Optimization Theory and Applications*, vol. 11, pp. 613–626, 1973.

[10] J. B. Cruz, "Leader-follower strategies for multilevel systems," *IEEE Transactions on Automatic Control*, vol. AC-23, pp. 244–255, 1978.

[11] T. Başar, "Information structures and equilibria in dynamic games," in *New Trends in Dynamic System Theory and Economics*, M. Aoki and A. Marzollo, Eds. New York and London: Academic Press, 1979, pp. 3–55.

[12] K. Okuguchi, *Expactations and Stability in Oligopoly Models*, ser. Lecture Notes in Economics and Mathematical Systems. Berlin: Springer-Verlag, 1976.

[13] R. Bellman, *Dynamic Programming*. Princeton, New Jersey: Princeton University Press, 1957.

[14] R. Isaacs, *Differential Games – A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. John Wiley & Sons, New York, 1965.

[15] T. Başar and H. Selbuz, "Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems," *IEEE Transactions on Automatic Control*, vol. AC-24, pp. 166–179, 1979.

[16] T. Başar, "Equilibrium strategies in dynamic games with multi-levels of hierarchy," *Automatica*, vol. 17, pp. 749–754, 1981.

[17] B. Tolwinski, "Closed-loop stackelberg solution to multi-stage linear-quadratic games," *Journal of Optimization Theory and Applications*, vol. 35, pp. 485–502, 1981.

[18] G. J. Olsder, "Phenomena in inverse Stackelberg problems," Mathematisches Forschungsinstitut Oberwolfach, Oberwolfach, Germany, Tech. Rep. 11/2005, 2005.

[19] K. Staňková, G. J. Olsder, and M. C. J. Bliemer, "Bi-level Optimal Toll Design Problem Solved by the Inverse Stackelberg Games Approach," in *Urban Transport XII: Urban Transport and the Environment in the 21st Century*, ser. The Built Environment, C. Brebbia and V. Doležel, Eds. Southampton, UK: WIT Press, 2006, vol. 89, pp. 871–880.

[20] ——, "Comparison of different toll policies in the dynamic second-best optimal toll design problem: Case study on a three-link network," *European Journal of Transport and Infrastructure Research*, vol. 9, no. 4, pp. 331–346, 2009.

[21] R. P. Hämäläinen, *Nash and Stackelberg solutions to general linear-quadratic two-player difference games*. Espoo, Finland: Systems Theory Laboratory B-29, Helsinki University of Technology publication, 1976.

[22] D. E. Kirk, *Optimal Control Theory, An Introduction*. Englewood Cliffs, New Jersey: Prentice Hall, 1970.

[23] D. P. Bertsekas, *Nonlinear Programming*. Belmont, Massachusetts: Athena Scientific, 2003.

[24] S. Boyd, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[25] A. Douglis, "The continuous dependence of generalized solutions of nonlinear partial differential equations upon initial data," *Communications on Pure and Applied Mathematics*, vol. 14, pp. 267–284, 1961.

[26] S. Kružkov, "First order quasilinear equations with several space variables," *Math. USSR-Sb.*, vol. 10, pp. 217–243, 1970.

[27] ——, "Generalized solution of the Hamilton-Jacobi equations of Eikonal type. i," *Math. USSR-Sb.*, vol. 27, pp. 406–446, 1975.

[28] R. P. Hämäläinen, *Generalized solutions of Hamilton-Jacobi equations*, ser. Pitman Research Notes Series. London: Pitman, 1982.

[29] M. G. Crandall and P.-L. Lions, "Viscosity Solutions of Hamilton-Jacobi Equations," *Transactions of the American Mathematical Society*, vol. 277, no. 1, pp. 1–42, 1983.

[30] M. Bardi and I. Capuzzo-Dolcetta, Eds., *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Boston: Birkhäuser, 2008.

[31] L. P. Lebedev and M. J. Cloud, *The Calculus of Variations and Functional Analysis With Optimal Control and Applications in Mechanics*. London: World Scientific Publishing Company, 2003.