Normal Form Proper Equilibria of Two-player Zero-sum Extensive Games

(extended abstract)

Peter Bro Miltersen and Troels Bjerre Sørensen

March 23, 2007

1 Introduction

A central solution concept in the theory of equilibrium refinements is the notion of *proper equilibria* of Myerson [3].

The normal form concept of properness can be applied to extensive games in two standard ways. An equilibrium in behavior strategies for an extensive form game with perfect recall is defined to be *normal form proper* if it is behaviorally equivalent to a proper equilibrium of the corresponding normal form game. This definition does not in any way restrict the behavior of a player in information sets *irrelevant* for the strategy of the player (i.e., information sets that the player is sure will not be reached if his strategy is played). In particular, it does not ensure subgame perfection. Van Damme [4] suggests a slightly more restrictive concept. He calls an equilibrium in behavior strategies of an extensive form game *induced* by a normal form proper equilibrium if it is a limit of behavior strategies induced by a sequence of ϵ-proper equilibria of the corresponding normal form. We shall adopt the slightly more convenient terminology *induced normal form proper equilibrium* for such an equilibrium. Van Damme showed that an induced normal form proper equilibrium is also quasi-perfect and hence sequential. It can be seen that the normal form proper equilibria are simply those equilibria that can be obtained by taking an induced normal form equilibrium and replacing the behavior in irrelevant information sets with arbitrary behavior. In particular, if we consider equilibria in behavior *plans* rather than behavior *strategies* the two notions coincide.

In this paper, we study normal form proper and induced normal form proper equilibria of *two-player zero-sum extensive form games with perfect recall.* We provide characterizations that enable these solution concepts to be computed efficiently (in theory as well as in practice) for a given game. In particular, we avoid the obvious approach of first converting the game to normal form (this obvious approach being inherently inefficient as it involves an exponential blowup in the size of the representation).

First, we study the case of *perfect information games* and show that the induced normal form proper equilibria of such games can be completely characterized by a certain *backwards induction* procedure, refining the standard backwards induction procedure for computing a subgame perfect equilibrium. The procedure can be easily implemented to run in linear time in the size of the game tree. As a curious example, applying the procedure to *tic-tac-toe* one finds that in any normal form proper equilibrium of this game, the game is opened by selecting the middle square with probability $\frac{1}{13}$. It may seem surprising that a proper equilibrium of the perfect information game tic-tac-toe cannot be pure. However, our characterization establishes that this fact is typical for combinatorial games (i.e., perfect information win/lose/draw games). Note that this is much unlike the case of perfect information games with generic payoffs where it is well known that there is a unique subgame perfect equilibrium and that this equilibrium is pure.

Second, we study the case of *imperfect information* games with perfect recall and show that the normal form proper equilibria of such games can be completely characterized by a procedure involving an iteratively defined sequence of linear programs derived from the linear programs for Nash equilibria in *sequence form* described by Koller, Megiddo and von Stengel [2]. Each linear program in the sequence has a number of variables and constraints which is at most the size of the game tree and the number of programs in the sequence is also at most the size of the game tree. The sequence of linear programs constructed is analogous to the linear programs arising in *Dresher's procedure* [1] established by van Damme [5] as characterizing the proper equilibria of a matrix game and the proof of correctness is based on relating our programs to the exponentially larger programs that would arise if Dresher's procedure was applied to the normal form of the game in consideration. As Koller, Megiddo and von Stengel, we represent equilibria by *realization plans*. Realization plans are in one-one correspondance with behavior plans. Hence, in the imperfect information case, we characterize the normal form proper equilibria, but not the induced ones.

2 Perfect information games

Let G be a perfect information zero-sum game played between Max, trying to maximize payoff and Min, trying to minimize payoff. The game is given by a game tree with payoffs in leaves and each internal node belonging to either Max, Min or Chance. For each node i in the tree we associate three number $\underline{v}_i \leq v_i \leq \overline{v}_i$. The number v_i is the usual minimax value of the node and may be computed by standard backwards induction. The values $\underline{v}_i$ and $\overline{v}_i$ can be informally seen as pessimistic and optimistic estimates of the expected outcome of the game from the point of Max, taking the possibility of mistakes being made by either player into account.

For a leaf with payoff p we let $\overline{v}_i = v_i = \underline{v}_i = p$. For an internal node i, we denote the set of immediate successors of i by $S(i)$ and define $\underline{v}_i, \overline{v}_i$ inductively as follows.

If i is a node belonging to Max, we let $V_i = (\cup_{j \in S(i)} \{v_j, \underline{v}_j\}) \setminus \{v_i\}$, i.e., the set of all values and all pessimistic estimates of all immediate successors of i, *except* the value of i itself. Then we let

$$\underline{v}_i = \begin{cases} \max(V_i) & \text{if } V_i \neq \emptyset, \\ v_i & \text{otherwise.} \end{cases} \tag{1}$$

Also, for a node i belonging to Max, we let $I_i = \{j \in S(i) | v_j = v_i \wedge \overline{v}_j > v_j\}$ and let

$$\overline{v}_i = v_i + \begin{cases} \frac{1}{\sum_{j \in I_i} (\overline{v}_j - v_j)^{-1}} & \text{if } I_i \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Similarly, if i is a node belonging to Min, we let $V_i = (\cup_{j \in S(i)} \{v_j, \overline{v}_j\}) \setminus \{v_i\}$, and let

$$\overline{v}_i = \begin{cases} \min(V_i) & \text{if } V_i \neq \emptyset, \\ v_i & \text{otherwise.} \end{cases} \tag{3}$$

Also, for node i belonging to Min, we let $I_i = \{j \in S(i) | v_j = v_i \wedge \underline{v}_j < v_j\}$ and let

$$\underline{v}_i = v_i - \begin{cases} \frac{1}{\sum_{j \in I_i} (v_j - \underline{v}_j)^{-1}} & \text{if } I_i \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

If i is a node belonging to Chance and $j \in S(i)$ is chosen by Chance with probability α_j , we let

$$\underline{v}_i = v_i - \min_{j \in S(i)} \alpha_j (v_j - \underline{v}_j) \tag{5}$$

$$\overline{v}_i = v_i + \min_{j \in S(i)} \alpha_j (\overline{v}_j - v_j) \tag{6}$$

In the full version of the paper we prove:

Theorem 1 *A behavior strategy profile ρ for G is an induced normal form proper equilibrium of G* if and only if *the following three conditions all hold:*

1. *For all nodes i and immediate successors j, ρ assigns non-zero behavior probability to j only if $v_i = v_j$.*
2. *For all nodes i belonging to Max for which $I_i \neq \emptyset$, ρ assigns behavior probability* exactly $\frac{(\overline{v}_j - v_j)^{-1}}{\sum_{j \in I_i} (\overline{v}_j - v_j)^{-1}}$ *to each $j \in I_i$.*
3. *For all nodes i belonging to Min for which $I_i \neq \emptyset$, ρ assigns behavior probability* exactly $\frac{(v_j - \underline{v}_j)^{-1}}{\sum_{j \in I_i} (v_j - \underline{v}_j)^{-1}}$ *to each $j \in I_i$.*

3 Imperfect information games

Let G be a two-player zero-sum extensive form game with perfect recall played between Max, trying to maximize payoff and Min, trying to minimize payoff. Our construction uses the *sequence form* of G and is based on the linear programming characterization of Nash equilibria in realization plans due to Koller, Megiddo and von Stengel [2]. We assume in this extended abstract that the reader is familiar with this paper and adopt its notation. In particular, G is given by a payoff matrix A, and realization plan constraint matrices E and F for Max and Min respectively. Let $v^{(0)}$ be the value of G. We define a series of linear programs (LPs) where the coefficients of each LP depend on the solutions of previous LPs. We group the LPs in pairs, denoting a pair as a round, the first being round 1. In round k, we first consider the LP (7). The vector variables x and q play the same roles as in the linear programs of Koller, Megiddo and von Stengel, except that they are now scaled by the scalar variable s. The vector variable u is indexed by *action sequences* of Min.

$$\begin{aligned} \max_{x,q,u,s} \quad & \mathbf{1}^\top u \\ \text{s.t.} \quad & -A^\top x + F^\top q + u + \sum_{0<i<k} m^{(i)} v^{(i)} s \le \mathbf{0} \\ & Ex - es = \mathbf{0} \\ & f^\top q - v^{(0)} s \ge 0 \\ & \mathbf{0} \le u \le \mathbf{1} \\ & x \ge \mathbf{0} \\ & s \ge 0 \end{aligned} \tag{7}$$

All optimal solutions to (7) agree on the value of the u-vector, which always takes the form of a 0/1-vector. Intuitively, the 1-entries in this 0/1-vector identify certain action sequences of Min as mistakes. Let $m^{(k)}$ be this optimal u. If $m^{(k)} \neq \mathbf{0}$, it defines (8):

$$\begin{aligned} \max_{x,q,t} \quad & t \\ \text{s.t.} \quad & -A^\top x + F^\top q + m^{(k)} t \le -\sum_{0<i<k} m^{(i)} v^{(i)} \\ & Ex = e \\ & f^\top q \ge v_0 \\ & x \ge \mathbf{0} \\ & t \ge 0 \end{aligned} \tag{8}$$

Here, the vector variables x and q are no longer scaled, and play the same roles as in the linear programs of Koller, Megiddo and von Stengel. The variable t is scalar. Let $v^{(k)}$ be the value of t in an optimal solution of (8). This completes a round. Informally, the x-part of an optimal solution to (8) is a realization plan for Max that is an optimal solution to the versions of (8) of all previous rounds, and among such optimal solutions optimally exploits the mistakes of Min defined by $m^{(k)}$.

In some round k, an optimal solution to (7) has $u = \mathbf{0}$ and the procedure is terminated. For this k, let D_1 be the set of optimal solutions to (8) in round $k-1$ (or, if $k = 0$, let D_1 be the set of maximin realization plans for Max). Interchanging the role of Max and Min and negating the payoff matrix, we carry out the entire procedure again. Let D_2 be the resulting set of realization plans for Min. In the full version of the paper, we prove:

Theorem 2 *$D_1 \times D_2$ is the set of normal form proper equilibria of G in realization plans.*

References

[1] M. Dresher. *The Mathematics of Games of Strategy: Theory and Applications*. Prentice-Hall, 1961.

[2] D. Koller, N. Megiddo, B. von Stengel. Fast algorithms for finding randomized strategies in game trees. In *Proc. 26th Ann. ACM Symposium on the Theory of Computing*, pages 750–759, 1994.

[3] R. B. Myerson. Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 15:133–154, 1978.

[4] E. van Damme. A relation between perfect equilibria in extensive form games and proper equilibria in normal form games. *International Journal of Game Theory*, 13:1–13, 1984.

[5] E. van Damme. *Stability and Perfection of Nash Equlibria*. Springer-Verlag, 2nd edition, 1991.