Computational complexity of correlated equilibria for extensive games

Bernhard von Stengel*

Department of Mathematics, London School of Economics
London WC2A 2AE, United Kingdom
stengel@maths.lse.ac.uk

CDAM Research Report LSE-CDAM-2001-03

August 9, 2001

Abstract. Correlated equilibria are more general than Nash equilibria. For games in strategic form, they are also easier to compute. This no longer holds when the game is given in extensive form. We show that for an extensive two-player game with perfect recall, even without chance moves, it is computationally difficult (NP-hard) to find a correlated equilibrium with maximum payoff sum. Even with a modified definition of correlated equilibria for extensive games, NP-hardness applies to any such concept that amounts to a distribution on pure strategy profiles.

Journal of Economic Literature Classification Number: C72

Keywords:
Correlated equilibrium, computational complexity, extensive games.

---

*Many thanks to Nimrod Megiddo, who suggested this problem, and to Françoise Forges, for stimulating discussions.

1. Introduction

Correlated equilibria are a generalization of Nash equilibria. One of their advantages is that they are simpler to compute. The set of correlated equilibria is a convex polyhedron, defined by the linear incentive constraints that express the equilibrium property. Finding a correlated equilibrium with maximum sum of the payoffs for all players amounts to maximizing a linear function over that polyhedron. The number of inequalities is polynomial in the size of the strategic form of the game (which is the input to the computational problem). The maximization problem can therefore be solved in polynomial time. Such computational problems are considered as "tractable" (see, for example, Papadimitriou, 1994). In contrast, computing a Nash equilibrium with maximum payoff sum is NP-hard ("computationally intractable" for larger problem size), even for two-player games (Gilboa and Zemel, 1989).

Computational advantages of correlated equilibria are no longer apparent when the game is given in extensive form. The reason is that the conversion to strategic form incurs an exponential blowup, since the number of strategies is typically exponential in the size of the game tree. However, if the game has two players which have perfect recall and the payoffs are *zero-sum*, then a Nash equilibrium can be found in polynomial time even in the size of the game tree. This has been first stated explicitly by Koller and Megiddo (1992). It follows also from a result of Romanovskii (1962), which until recently has been overlooked in the English-speaking community. A related proof uses the *sequence form* of an extensive game with perfect recall, due to von Stengel (1996). This is a strategic description of the game which is as useful as the reduced strategic form of the game, but which has proportional rather than exponential size compared to the game tree.

Could the sequence form, or a related strategic representation, be used to compute correlated equilibria of an extensive game? Compared to Nash equilibria, this raises the following conceptual difficulty. A correlated equilibrium is in effect played such that a "mediator" recommends a pure strategy to each player. From the received recommendation, each player has a posterior distribution on the recommendations given to the other players. The own recommended strategy must be a best response; this defines the incentive constraints. In a Nash equilibrium, that posterior distribution is always the same, but in a correlated equilibrium it may vary with each own pure strategy. Describing these distributions for each own strategy seems to require an exponential number of variables.

This suggests that strategic-form correlated equilibria of extensive games may be hard to compute. Indeed, this is a result of the present paper.

An alternative definition of correlated equilibria for extensive games with perfect recall may be closer in spirit to *behavior strategies*. Behavior strategies suffice for Nash equilibria, and are represented in the sequence form. Forges and von Stengel (2001) propose the "extensive form correlated equilibrium" as a such concept. In essence, it defines correlated recommendations of *moves* at information sets as these are reached, rather than recommendations of strategies at the beginning of the game. This captures a larger set of equilibria, for example as discussed by Forges (1986, 1993) and Myerson (1986, 1994). However, even this extensive form correlated equilibrium amounts to a delayed recommendation of a *pure strategy* to each player. The NP-hardness result below therefore applies to this concept as well.

Our result states that any concept of equilibrium for extensive games that results in a distribution on pure strategy profiles, and includes all Nash equilibria, gives rise (via payoff maximization) to an NP-hard computational problem. The proof uses a rather simple reduction from the satisfiability problem. A logical formula is translated to an extensive game that has a strategic form similar to the truth table for the formula. The players have identical payoffs. The maximum payoff of an equilibrium shows if the formula is satisfiable or not. A satisfying assignment corresponds to certain pair of pure strategies of the two players.

In order to avoid chance moves, we also present a construction where an "outside option" in the form of zero-sum game forces one of the players to randomize in equilibrium. This shows that for two-player games without chance moves, finding a correlated equilibrium for the strategic form with maximum payoff sum is NP-hard. The incentive constraints are used to prove this result. Unlike the previous construction with chance moves, this result does not apply to other definitions of correlated equilibria for extensive games. However, it applies to games with three players.

As mentioned, Gilboa and Zemel (1989) showed that finding a Nash equilibrium with maximum payoff sum in a bimatrix game is NP-hard. Koller and Megiddo (1992) proved that it is NP-hard to find a max-min behavior strategy for a player without perfect recall in an extensive two-person zero-sum game. Blair, Mutchler, and van Lent (1996, Section 4.2) showed that it is NP-hard to decide if an extensive two-person zero-sum game with imperfect information but perfect recall has an equilibrium in pure strategies. Frank and Basin (2001) show the NP-hardness of a concept of "best defense" in games with imperfect information. Even for bimatrix games, the computational complexity of *finding one* Nash equilibrium — not necessarily with maximum payoff sum — is still unclear (see Megiddo, 1988, and Megiddo and Papadimitriou, 1989).

2. Basic definitions

We use the standard definition of games in extensive form. A game tree with information sets describes the players, their choices, and their payoffs at the leaves (terminal nodes) of the tree. Every player has *perfect recall*, that is, any two nodes in an information set of the player are preceded by the same sequence of earlier moves of that player. A *strategy* defines a move for every information set of a player. In the *reduced strategic form*, strategies are identified that differ only in moves at information sets which are unreachable due to an earlier own move of the player. Without loss of generality, Nash and correlated equilibria can be considered directly for the reduced strategic form.

A *correlated equilibrium* of a two-player game can be defined as follows (see also Myerson, 1994). Let i and j stand for strategies of player I and II, respectively, with resulting payoffs a_{ij} and b_{ij}. A correlated equilibrium is a distribution on strategy pairs. When a strategy pair (i,j) is drawn according to this distribution, player I is told i and player II is told j. The probabilities x_{ij} are nonnegative and sum up to one and must fulfill the following *incentive constraints*. For all strategies i and k of player I and all strategies j and l of player II,

$$\begin{aligned} \sum_j a_{ij}\, x_{ij} &\ge \sum_j a_{kj}\, x_{ij}\,, \\ \sum_i x_{ij}\, b_{ij} &\ge \sum_i x_{ij}\, b_{il}\,. \end{aligned} \tag{1}$$

These inequalities state that player I, when recommended to play i, has no incentive to switch from i to k, given (up to normalization) the conditional probabilities x_{ij} on opponent strategies j. Similarly, the second inequalities in (1) state that player II, when recommended to play j, has no incentive to switch to l.

We recall the following notions from computational complexity (see Papadimitriou, 1994). A *decision problem* consists of a set of *instances* which are strings of text that encode an input to the problem, for example a graph. The computational problem is then to decide if that instance fulfills a certain property, for example if the graph has a Hamiltonian cycle. The *satisfiability problem* SAT has as an instance a Boolean formula ϕ in conjunctive normal form. That is, ϕ is a conjunction of *clauses*, each of which is a disjunction of *literals,* each of which is an unnegated or negated Boolean variable. The question is if ϕ is *satisfiable,* that is, has an assignment of truth-values to its variables so that the formula evaluates to true.

A *polynomial (-time) algorithm* terminates on all input strings in a number of steps that is bounded by a polynomial in the length of the string. A decision problem is *NP-complete* if there is a polynomial algorithm that transforms each instance ϕ of SAT to an instance of the problem, so that the instance has the property in question if and only if ϕ is satisfiable. (The problem must also belong to NP, which means that whenever an instance has the property, this can be verified in polynomial time with help of a short "certificate", for example the satisfying truth assignment for a SAT instance.) It is widely believed that $P \neq NP$, that is, there is no polynomial algorithm for solving an NP-complete problem. The best known algorithms for SAT are exponential, essentially no better than trying out all truth assignments.

An *optimization problem* is like a decision problem except that any instance has a *value*, which is a number rather than just the answer "yes" or "no". An optimization problem is called *NP-hard* if, for instances of that problem and arbitrary numbers k, the decision problem whether the value of the instance is at least k is NP-complete.

3. Equilibria in games with chance moves

Consider a game in extensive form. A small description, by a polynomial number of inequalities, of the set of correlated equilibria would imply that a linear function over that set, for example the sum of payoffs to the players, can be maximized in polynomial time. If the latter problem is NP-hard, then no small description exists, unless $P = NP$.

The strategic form is in general exponentially larger than the extensive form. It is therefore plausible that correlated equilibria as defined for the strategic form have no small description. For extensive games, other definitions of correlated equilibria are known (for details see Forges, 1986, 1993; Myerson, 1986; Forges and von Stengel, 2001). To some extent, they consider moves rather than strategies, which may offer a potential reduction in size, as the sequence form does for Nash equilibria (see von Stengel, 1996).

Most of these definitions share the following property. At the beginning of the game, a mediator selects a move for every information set according to a known distribution. Every player is told his moves either at the beginning of the game (as a pure strategy), or at the beginning of a *stage* in a multi-stage game, or, as in the *agent normal form*, when the information set is reached. In effect, this amounts to a *pure strategy* recommended to each player, even though the moves in that strategy may not all be told in advance. However, we show that all such concepts of "equilibria" (which we merely assume to include all Nash equilibria) lead to an NP-hard optimization problem. Our result applies also to strategic-form correlated equilibria, and to *distribution equilibria* (see Sorin, 1998).

Theorem 1. *Consider for any two-player extensive game with perfect recall a set of "equilibria", which are convex combinations of pure strategy pairs, that includes all Nash equilibria. Then the problem of finding an equilibrium with maximum payoff sum is NP-hard.*

Proof. We give a polynomial transformation of SAT to the decision problem "does the game have an equilibrium with payoff sum at least two", which shows that finding an equilibrium with maximum payoff sum is NP-hard. Consider a Boolean formula ϕ in conjunctive normal form with n clauses and m variables. The question is whether ϕ is satisfiable.

We construct a two-player game from ϕ as follows; Figure 1 shows an example. An initial chance move chooses with probability $1/n$ one of the n decision points of player II, which correspond to the clauses. Player II is fully informed about the chance move and, for each clause chosen, selects one of the literals in the clause, representing the literal that is to be true in the clause (which exists for each clause if and only if ϕ is satisfiable). The respective move of player II leads to an information set of player I given by the variable in the chosen literal.

Player I has m information sets, corresponding to the m variables, and has two choices at each information set, representing the two possible Boolean values for that variable. Player I receives no information about the move of player II, and is therefore ignorant about the chance move and whether a variable is chosen to make a particular clause true. After player I's move, the game terminates and the players receive an identical payoff, which is equal to one if the chosen clause (by chance), literal (by player II) and truth-value for the literal (by player I, given by the truth value for the variable) is true,

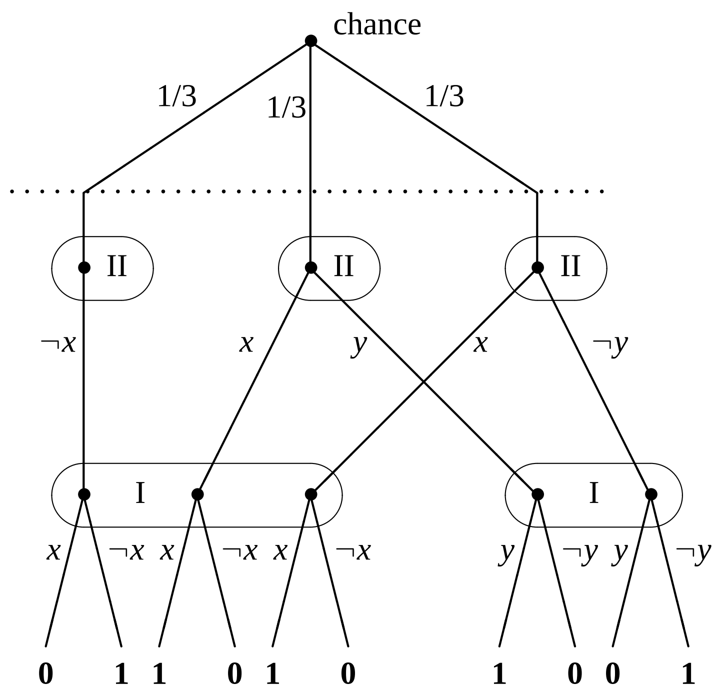

Figure 1. Extensive game for the SAT instance $(\neg x) \wedge (x \vee y) \wedge (x \vee \neg y)$, which is not satisfiable. Chance chooses a clause (this part, above the dotted line, is replaced in Figure 2). Player II picks a literal within the clause, and player I for each variable the literal to be made true. The payoffs **0** and **1** are the same for both players.

and zero otherwise. The 2^m pure strategies of player I are therefore the possible truth assignments to the variables in ϕ.

In this game, there is a pair of pure strategies for the two players with payoff one if and only if ϕ is satisfiable: Clearly, if ϕ is satisfiable, then the satisfying assignment defines a pure strategy for player I, and player II can pick for each clause a literal that makes the clause true, which defines a pure strategy for player II. Then each possible move of nature leads to a payoff of one, which is the overall payoff. Conversely, if ϕ is not satisfiable, then any truth assignment to the m variables necessarily has at least one clause that is false, so that the respective move of player II will lead to a payoff zero. The overall expected payoff is then at most $1 - 1/n$. In that case, any convex combination of pure strategy pairs has also at most payoff $1 - 1/n$ per player.

A pure strategy pair with payoff one for each player defines a Nash equilibrium since that is the maximum possible payoff. Hence, an equilibrium with maximum payoff sum shows if ϕ is satisfiable, so computing such an equilibrium is an NP-hard problem. $\square$

4. Equilibria in games without chance moves

The construction in the proof of Theorem 1 uses an initial chance move. In consequence, both players have an exponential number of pure strategies. The strategies of player I represent truth assignments to Boolean variables, like the rows in a truth table for the formula. Player II's strategies select independently for each clause a literal to be made true. For an instance of 3SAT, which has exactly three literals in each of the n clauses, player II then has 3^n strategies.

The strategic form of the game is large even if only one player has an exponential number of strategies. Our second result shows that even for two-player games without chance moves, it is NP-hard to find a strategic-form correlated equilibrium with maximum payoff sum.

Such an NP-hardness result cannot hold when the maximum payoff sum applies to an equilibrium in pure strategies, as in the construction for Theorem 1. Without chance moves, any pair of pure strategies leads to a leaf of the game tree, so one could simply inspect all leaves to find the maximum payoff sum. Hence, the players must somehow be forced to randomize.

Furthermore, it is — at least in the approach taken below — no longer possible to make a strong statement about arbitrary convex combinations of pure strategy pairs as in Theorem 1. Even if the maximization problem for Nash equilibria is NP-hard, suitable definitions of correlated equilibria for the extensive form may lead to a distribution on leaves with larger payoffs sums. Our construction uses the definition of correlated equilibrium for the strategic form.

The proof extends that of Theorem 1. Player II replaces chance, and in equilibrium assigns positive probability to every clause. Otherwise, player I uses one of n initial *outside options*, which is not in player II's interest.

Theorem 2. *For extensive two-player games with perfect recall and without chance moves, it is NP-hard to find a strategic-form correlated equilibrium with maximum payoff sum.*

Proof. As in the proof of Theorem 1, consider again a Boolean formula ϕ in conjunctive normal form with n clauses. In the game constructed above, chance can only be replaced by player II since if it was player I, that player would not have perfect recall. So, let player II move first, with n moves $c_1, \ldots, c_n$ corresponding to the clauses of ϕ. We then introduce an extra information set for player I who is uninformed of the choice of player II. There, player I has $n+1$ choices $o_1, \ldots, o_n$ called *outside options* which terminate the game, and an extra choice *in* of staying in the game, which then continues as before. Figure 2 gives an illustration.

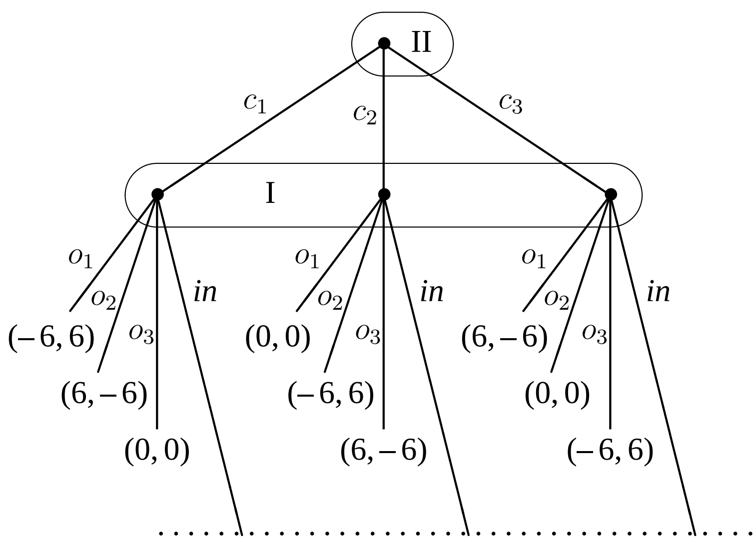

Figure 2. Pre-play of a zero-sum game between player II and player I that induces player II to completely mix between her choices $c_1, \ldots, c_n$. Otherwise player I chooses among his outside options $o_1, \ldots, o_n$ instead of *in*. After the dotted line, the game continues as in Figure 1.

What is the reduced strategic form of this game? Player II has exactly as many reduced strategies as ϕ has literals, since a strategy consists of a choice of a clause and a literal therein; the remaining choices of literals are irrelevant. Player I's strategies are the truth assignments, each preceded by the move *in*, and n extra strategies given by the initial outside options $o_1, \ldots, o_n$.

The n outside option rows of player I and the n columns $c_1, \ldots, c_n$ for the clauses (each replicated with its number of literals, which we ignore for the moment) form a zero-sum game with a unique, completely mixed equilibrium and value zero. The payoffs in this zero-sum game are scaled such that whenever player II does not choose a clause, player I can get a payoff of at least 2, and thus will not choose *in*, which is, however, not part of an equilibrium.

The zero-sum payoffs are a variation of "rock–scissors–paper". Suppose that when player I chooses i and player II chooses j in $\{1, \ldots, n\}$, player I gets from player II the amount

$$a_{ij} = (j - i) \bmod n\,.$$

The matrix of these payoffs for $n = 5$, for example, is

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 & 3 \\ 3 & 4 & 0 & 1 & 2 \\ 2 & 3 & 4 & 0 & 1 \\ 1 & 2 & 3 & 4 & 0 \end{bmatrix}. \tag{2}$$

Clearly, the uniform mixed strategy to play each row (respectively, column) with probability $1/n$ is optimal for both players. The value v of this game is $(n-1)/2$.

Suppose that player II chooses one of the columns with probability $q < 1/n$. By symmetry, let this be column 1. Let player I respond by playing row 1 (with the bad payoff 0 for the under-played column) with probability $2/n$, row 2 (with the good, but now under-used payoff $n-1$) with probability zero, and all other rows as before with probability $1/n$. The resulting payoffs to player II are easily found by comparison with the uniform mixed strategy, since only two rows are played differently. In column 1, the expectation is $v - (n-1)/n$ (where the probability $1/n$ is shifted from payoff $n-1$ to payoff 0), and in each other column $j = 2, \ldots, n$ it is $v + 1/n$ since there we have $a_{1j} = a_{2j} + 1$, as (2) demonstrates. The latter columns have combined probability $1 - q$. The resulting payoff to player I is therefore

$$(v - (n-1)/n)\,q + (v + 1/n)(1-q),$$

or $v + 1/n - q$. Player II therefore has to pay an extra amount of at least $1/n - q$ compared to playing all columns uniformly and having to pay only v.

We re-scale the payoffs a_{ij} by subtracting v, so that the game has value zero, and then multiply by $2n$. Then if player II does not play a column at all ($q = 0$), player I can gain at least 2. If the smallest probability q for a column is less than $1/2n$, then player I will gain more than 1.

With this re-scaling, any leaf of the game tree following move c_j by player II and move o_i by player I gets the zero-sum payoff

$$2n((j-i) \bmod n) - n(n-1)$$

to player I. The preceding discussion shows that to make player I choose *in*, player II must assign at least probability $1/2n$ to every clause. Furthermore, choosing each clause with probability $1/n$ will make player I stay in, since then any outside option has expected payoff zero, but some clauses can be made true, which gives positive payoff.

If ϕ is satisfiable, then there is a correlated equilibrium with payoff one to each player and payoff sum 2, as before. Suppose ϕ is not satisfiable, and consider a correlated equilibrium. The maximum payoff sum for any pure strategy pair is 2, where the pure strategy of player I is not an outside option. Consider such a truth assignment that is recommended to player I with positive probability (outside options will only contribute zero to the expected payoff sum). Then some entries in that row of the strategic form have payoff zero, and have probability at least $1/2n$, since otherwise player I would switch to an outside option. Thus, the conditional expected payoff to player I is at most $1 - 1/2n$, and the sum at most $2 - 1/n$. Therefore, the maximum payoff sum is at most $2 - 1/n$. Finding a correlated equilibrium with maximum payoff sum would therefore decide if ϕ is satisfiable. $\square$

The preceding proof is not suitable for a definition of correlated equilibrium where the players receive *delayed* recommendations. Consider Figure 2, and let a mediator decide with probability 1/3 between the following recommendations to the players. Player I is always told to choose *in* and y. Player II is told with equal probability to choose c_1, c_2, or c_3, and the literal containing the variable x. These choices are correlated with the recommendations $\neg x$, x, and x to player I, respectively. These result in payoff 1 to both players, despite the fact that the given formula is not satisfiable.

In the strategic form, this does not define an equilibrium of the game since player I can exclude at least one clause when being told his full strategy, and would switch to an outside option. However, when player I is *only* told to choose *in* at his first information set, awaiting further instructions later, then each of c_1, c_2, c_3 has probability 1/3 and player I will follow the recommendation. This is similar to a classic example by Myerson (1986) showing that delayed recommendations can lead to higher payoffs.

Could correlated equilibrium concepts involving delayed recommendations be easy to compute for games without chance moves? For games with three players, the answer is again negative, as our final discussion shows.

Consider the construction in Theorem 1 and replace chance by a third player III. At a leaf of the game tree where player I and II receive zero, so does III, and when they receive payoff 1, player III gets -1. From the perspective of player III, this is a zero-sum game where III has to pay one unit if ϕ is satisfiable, but pays at most $1 - 1/n$ if not, by choosing every clause with equal probability $1/n$.

Any leaf with payoffs $(1, 1, -1)$ to the three players has payoff sum one, so the maximum payoff sum in any convex combination of pure strategies is one. Clearly, if ϕ is satisfiable, there is a Nash equilibrium with payoff sum one, as in Theorem 1.

Suppose that ϕ is not satisfiable. Unlike in Theorem 1, there are pure strategy combinations with payoff sum one, but they do not form an equilibrium. For that, we make the following assumption on what constitutes an equilibrium: When player III gets a recommendation, the recommendations to the other players are given by some probability distribution which may depend on what player III has been told, but not on what player III actually does. That is, player III can expect a definite behavior of player I and II when considering whether to follow the recommendation or not. This seems basic for a definition of equilibrium. Player III moves only once, so delays are not an issue.

Once player III has been given a recommendation, the behavior of player I and II is some distribution on pure strategy pairs for the two players. None of these has all clauses true, for otherwise ϕ would be satisfiable. Hence, *some* move of player III leads with positive probability to a leaf with payoff zero. If this is the recommended move, the expected payoff sum is less than one. If the recommended move has expected payoff sum one, player III will deviate, which violates the equilibrium condition. Thus, as before, if ϕ is not satisfiable, no equilibrium has the maximum payoff sum one.

This shows that the set of correlated equilibria (fulfilling the above assumption) cannot have a small description for extensive three-player games without chance moves.

Since the set of correlated equilibria should be a polytope for any number of players, it seems unlikely that the two-player case is significantly different. At present, the question for two players remains open.

References

J. R. S. Blair, D. Mutchler, and M. van Lent (1996), Perfect recall and pruning in games with imperfect information. Computational Intelligence 12, 131–154.

F. Forges (1986), An approach to communication equilibria. Econometrica 54, 1375–1385.

F. Forges (1993), Five legitimate definitions of correlated equilibrium in games with incomplete information. Theory and Decision 35, 277–310.

F. Forges and B. von Stengel (2001), Extensive form correlated equilibria. Manuscript, London School of Economics.

I. Frank and D. Basin (2001), A theoretical and empirical investigation of search in imperfect information games. Theoretical Computer Science 252, 217–256.

I. Gilboa and E. Zemel (1989), Nash and correlated equilibria: some complexity considerations. Games and Economic Behavior 1, 80–93.

D. Koller and N. Megiddo (1992), The complexity of two-person zero-sum games in extensive form. Games and Economic Behavior 4, 528–552.

N. Megiddo (1988), A note on the complexity of p-matrix LCP and computing an equilibrium. Research Report RJ 6439, IBM Almaden Research Center, San Jose, California.

N. Megiddo and C. H. Papadimitriou (1989), On total functions, existence theorems and computational complexity (Note). Theoretical Computer Science 81, 317–324.

R. B. Myerson (1986), Multistage games with communication. Econometrica 54, 323–358.

R. B. Myerson (1994), Communication, correlated equilibria and incentive compatibility. In: Handbook of Game Theory, Vol. II, eds. R. J. Aumann and S. Hart, Elsevier, Amsterdam, 827–847.

C. H. Papadimitriou (1994), Computational Complexity. Addison-Wesley, Reading, Mass.

I. V. Romanovskii (1962), Reduction of a game with complete memory to a matrix game. Soviet Mathematics 3, 678–681.

S. Sorin (1998), Distribution equilibrium I: definition and examples. Cahier du THEMA 9835, Paris.

B. von Stengel (1996), Efficient computation of behavior strategies. Games and Economic Behavior 14, 220–246.